

**FORSCHUNGSZENTRUM JÜLICH GmbH**  
**Zentralinstitut für Angewandte Mathematik**  
**D-52425 Jülich, Tel. (02461) 61-6402**

Interner Bericht

**Performance-Untersuchung zum Einfluß  
der Netzwerkinfrastruktur auf  
NFS-Anwendungen in Workstation-Gruppen**

*Werner Anrath, Ralf Malzkorn,  
Olaf Mextorf, Sabine Werner*

KFA-ZAM-IB-9607

März 1996  
(Stand 01.03.1996)

<b>EINFÜHRUNG .....</b>	<b>4</b>
<b>TESTUMGEBUNG .....</b>	<b>5</b>
Systemkonfiguration .....	5
Testoperation 1: Lesen über NFS .....	6
Testoperation 2: Suchen über NFS .....	7
Testoperation 3: Schreiben über NFS.....	7
<b>TESTERGEBNISSE .....</b>	<b>8</b>
<b>1. Gemeinsames Ethernet für die Workstation-Gruppe und ca. 200 weitere Netzwerkteilnehmer .....</b>	<b>8</b>
1. a. Leseoperation bei einem gemeinsamen Ethernet für die Workstation-Gruppe und weitere Netzwerk- teilnehmer.....	9
1. b. Suchoperation bei einem gemeinsamen Ethernet für die Workstation-Gruppe und weitere Netzwerk- teilnehmer.....	9
1. c. Schreiboperation bei einem gemeinsamen Ethernet für die Workstation-Gruppe und weitere Netz- werkteilnehmer .....	10
1. d. Zusammenfassung der Ergebnisse zur Netzwerkkonfiguration 1 .....	10
<b>2. Separates Ethernet für die Workstation-Gruppe.....</b>	<b>11</b>
2. a. Leseoperation bei einem separaten Ethernet für die Workstation-Gruppe .....	12
2. b. Suchoperation bei einem separaten Ethernet für die Workstation-Gruppe .....	13
2. c. Schreiboperation bei einem separaten Ethernet für die Workstation-Gruppe.....	14
2. d. Zusammenfassung der Ergebnisse zur Netzwerkkonfiguration 2 .....	15
2. e. Operationen parallel .....	15
<b>3. Einsatz eines Ethernet-FDDI-Switches zwischen Gruppenserver und Clients.....</b>	<b>17</b>
3. a. Schreiboperation beim Einsatz eines Ethernet-FDDI-Switches .....	18
3. b. Suchoperation beim Einsatz eines Ethernet-FDDI-Switches.....	19
3. c. Schreiboperation beim Einsatz eines Ethernet-FDDI-Switches .....	20
3. d. TCPSPRAY beim Einsatz eines Ethernet-FDDI-Switches.....	21
3. e. Zusammenfassung der Ergebnisse zur Netzwerkkonfiguration 3 .....	22
<b>4. Anbindung eines Clients über FDDI an den Server .....</b>	<b>23</b>
4. a. Performance-Werte beim Einsatz eines FDDI-Ringes zwischen Client und Server .....	23
4. b. Zusammenfassung der Ergebnisse zur Netzwerkkonfiguration 4 .....	24
<b>PREISVERGLEICH.....</b>	<b>25</b>
<b>HINTERGRUNDINFORMATIONEN ZU NFS.....</b>	<b>26</b>
NFS-RPCs und Datenvolumen während der Tests im Vergleich.....	27
Einfluß des Caches auf NFS-Operationen .....	28
NFS-Statistiken aus dem Produktionsbetrieb.....	29
<b>HINTERGRUNDINFORMATIONEN ZU SCSI.....</b>	<b>31</b>
<b>ZUSAMMENFASSUNG .....</b>	<b>33</b>
<b>LITERATUR.....</b>	<b>35</b>

## ABBILDUNGEN

Abbildung 1: die Workstationkonfiguration während der Tests.....	5
Abbildung 2: Netzwerkkonfiguration 1 .....	8
Abbildung 3: Ethernet-Statistik während des Tests 1. a. ....	9
Abbildung 4: Ethernet-Statistik während des Tests 1. b. ....	9
Abbildung 5: Ethernet-Statistik während des Tests 1. c. ....	10
Abbildung 6: Netzwerkkonfiguration 2 .....	11
Abbildung 7: Ethernetstatistik des Tests 2. a. ....	12
Abbildung 8: Ethernetstatistik des Tests 2. b. ....	13
Abbildung 9: Ethernet-Statistik des Tests 2. c. ....	14
Abbildung 10: Antwortzeiten bei paralleler und sequentieller Ausführung der Tests .....	15
Abbildung 11: Antwortzeiten aller Tests mit der Netzwerkkonfiguration 2. ....	16
Abbildung 12: Netzwerkkonfiguration 3 .....	17
Abbildung 13: FDDI-Statistik zum Test 3. a. ....	18
Abbildung 14: Ethernet-Statistik des ersten Switchports zum Test 3. a. ....	18
Abbildung 15: FDDI-Statistik zum Test 3. b. ....	19
Abbildung 16: Ethernet-Statistik des ersten Switchports zum Test 3. b. ....	19
Abbildung 17: FDDI-Statistik zum Test 3. c. ....	20
Abbildung 18: Ethernet-Statistik des ersten Switchports zum Test 3. c. ....	20
Abbildung 19: FDDI-Statistik zum Test 3. d. ....	21
Abbildung 20: Ethernet-Statistik des ersten Switchports zum Test 3. d. ....	21
Abbildung 21: Antwortzeiten aller Tests mit der Netzwerkkonfiguration 3. ....	22
Abbildung 22: Netzwerkkonfiguration 4 .....	23
Abbildung 23: Anzahl der NFS-RPCs der Testoperationen im Vergleich .....	27
Abbildung 24: bewegte Datenvolumen der Testoperationen im Vergleich .....	27
Abbildung 25: NFS-Server-Statistiken in AXP-Workstation-Gruppen.....	29
Abbildung 26: NFS-Server-Statistiken in AIX-Workstation-Gruppen.....	29
Abbildung 27: Vergleich der Anzahl der RPC-Requests einer Client-Workstation in den verschiedenen Workstation-Gruppen .....	30

## TABELLEN

Tabelle 1: Aktivierungszeiten einiger typischer Applikationen .....	6
Tabelle 2: Performance-Werte auf den Workstations zum Test 2. a. ....	12
Tabelle 3: Ethernet-Statistik des Gruppenservers zum Test 2. a. ....	12
Tabelle 4: Performance-Werte auf den Workstations zum Test 2. b. ....	13
Tabelle 5: Performance-Werte auf den Workstations zum Test 2. c. ....	14
Tabelle 6: Ethernet-Statistik des Gruppenservers zum Test 2. c. ....	14
Tabelle 7: Performance-Werte auf den Workstations zum Test 3. a. ....	18
Tabelle 8: Performance-Werte auf den Workstations zum Test 3. b. ....	19
Tabelle 9: Performance-Werte auf den Workstations zum Test 3. c. ....	20
Tabelle 10: Performance-Werte auf den Workstations zum Test 3. d. ....	21
Tabelle 11: Performance-Werte auf den Workstations zum Test 4 .....	23
Tabelle 12: Kosten der vorgestellten Netzwerkkonfigurationen .....	25
Tabelle 13: SCSI II - Bus - Spezifikationen .....	31
Tabelle 14: Leistungsdaten aktueller SCSI-Laufwerke (Herstellerangaben).....	32

## Einführung

Das ZAM unterstützt seit rund zwei Jahren in der KFA den Aufbau von sogenannten Workstation-Gruppen, die aus einer Anzahl von Client-Systemen und einem File-Server bestehen. Der File-Server übernimmt in dieser Konfiguration vorzugsweise die Speicherung und Bereitstellung von Benutzerdaten und Applikationen über NFS (Network File System).

Bisher sind nur in begrenztem Umfang bei der Netzerkennung solcher Arbeitsgruppen Aspekte wie Lastentkopplung und Netzwerk-Performance berücksichtigt worden. Die Erfahrungen aus dem Betrieb der Workstation-Gruppen zeigen bei stetig wachsender CPU- und Grafikleistung der Client-Workstations stark unterschiedliche Leistungsanforderungen bezogen auf den jeweiligen File-Server und die Netzwerkkomponenten. Diese Leistungsanforderungen, erwachsen u.a. aus dem Leistungsbedarf heutiger offener Betriebssysteme, sind hochgradig applikationsabhängig. Welche Anzahl von Client-Systemen an einem File-Server unter Ausnutzung einer bestimmten Netzwerktechnologie sinnvoll betrieben werden kann, läßt sich im voraus nicht festlegen. Daher verlangt insbesondere die Auswahl der File-Server-Hardware sowie der aktiven Netzwerkkomponenten anpassungsfähige und skalierbare Lösungen.

Bei den angebotenen Massenspeichern, i.a. SCSI-Geräte, ist derzeit unter Ausnutzung optionaler SCSI II Erweiterungen wie Fast SCSI und Wide SCSI noch am ehesten eine kontinuierliche Leistungssteigerung (nach einer langen SCSI I Durststrecke) zu beobachten. Die im folgenden präsentierten Messungen und Testumgebungen legen einerseits die Leistungsgrenzen in bestimmten Umgebungen offen und zeigen andererseits stark wechselnde Anforderungen an die gemeinsam genutzten Komponenten wie Netzwerk und File-Server.

# Testumgebung

## Systemkonfiguration

Die Tests wurden mit einer auf der DEC-AXP-Familie basierenden Workstation-Gruppe durchgeführt. Die entsprechende Systemkonfiguration, die bei jedem Test gleich blieb, ist folgendem Diagramm zu entnehmen:

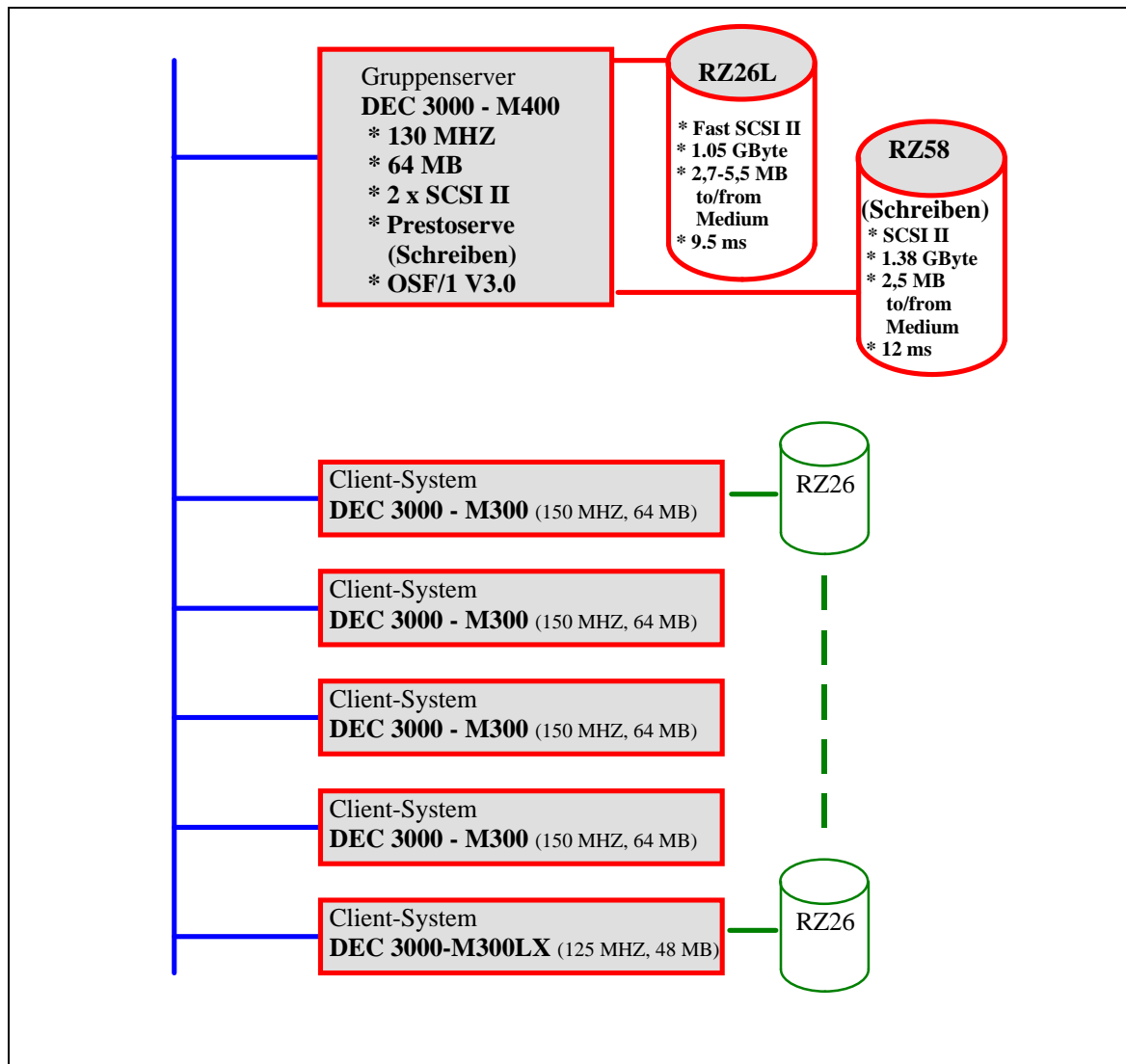


Abbildung 1: die Workstationkonfiguration während der Tests

Bei den Tests wurden gleiche, teils das Netzwerk stark belastende NFS-Operationen in unterschiedlichen Netzwerkkonfigurationen durchgeführt, die den einzelnen Clients und dem Server unterschiedliche Kommunikationsbandbreiten zur Verfügung stellten.

Diese Netzwerkkonfigurationen variierten von einem Anschluß der gesamten Workstation-Gruppe in einem flachen Ethernet mit einer großen Anzahl weiterer Netzwerkteilnehmer bis zum Anschluß des Servers und eines Clients an einem exklusiven FDDI-Ring.

In jeder Netzwerkkonfiguration wurden Lese-, Such- und Schreiboperationen über NFS zwischen dem Server und den Clients ausgeführt. Jede Operation wurde dabei nacheinander von einem bis zu fünf Clients gleichzeitig durchgeführt, um das Antwortzeitverhalten in Abhängigkeit von der Anzahl der Operationen (Clients gleichzeitig) vergleichen zu können.

Im folgenden werden zunächst die NFS-Operationen und deren Hintergrund vorgestellt und anschließend die Testergebnisse zu den unterschiedlichen Netzwerkkonfigurationen präsentiert, wobei sowohl die auf den Testrechnern ermittelten Ausführungszeiten, CPU-Lasten und Transferraten als auch die

direkt mit Hilfe der involvierten Netzwerkkomponenten auf dem Netz gemessenen Statistiken aufgeführt werden.

Bei der Ausführung der Tests werden auf den Workstations sowohl die real benötigte Zeit als auch die NFS- und ETHERNET-Statistik protokolliert (*nfsstat*, *netstat*). Auf der File-Server-Seite werden I/O-Operationen und CPU-Belastung (*iostat*) aufgezeichnet. Diese Daten sollen möglichst präzise die Wechselwirkung zwischen Applikation (Lesen), Netzwerk und File-Server-Leistung verdeutlichen.

Ausgewertet werden nur Daten, bei denen während der Tests keine anderen nennenswerten Aktivitäten (z.B. File-Server-Backup mit Datenkompression) nachweisbar waren; da die Workstation-Gruppe jedoch insgesamt im Normalbetrieb (Zeitsynchronisation, Mail, Netzwerk-Überwachung....) lief, müssen minimale Streuungen in Kauf genommen werden.

Beispiel: Empfängt der File-Server während eines Tests einige *Mails*, so ergibt sich eine CPU-Last von 1-2% mehr.

Eine Backup-Operation mit Datenkompression auf dem File-Server würde die CPU-Last auf 50-80% steigern, die NFS-bedingte CPU-Last kann dann aus der *iostat*-Statistik nicht abgelesen werden.

### **Testoperation 1: Lesen über NFS**

Bei diesem Test lesen eine bis fünf Client-Workstations jeweils verschiedene 50 MByte große Dateien, die auf der RZ26L des Servers gespeichert sind, über NFS (NFS V2 read) vom Server.

Dabei führen die Client-Systeme crontab-gesteuert "*cp file\_50MB.dat /dev/null*" aus. Die Systemzeiten dazu werden durch das XNTPD-Protokoll synchronisiert.

Der File-Server wird durch diesen Test von jeder Workstation mit mehr als 6000 RPCs (NFS-Remote-Procedure Calls) belastet. Der gleichzeitige andauernde Zugriff auf verschiedene Dateien sorgt für eine angemessene Beanspruchung des I/O-Subsystems (Cache, Treiber, Host-Adapter, SCSI-BUS, Platte). Leseanforderungen über NFS können durch die Verwendung von Dateisystem-Caches sowohl auf der Client- als auch auf der File-Server-Seite effizient abgearbeitet werden. Durch Anpassen bestimmter *Mount*-Parameter und Starten zusätzlicher *NFSD* - und *NFSIOD*-Prozesse kann die Leistungsfähigkeit begrenzt verbessert werden. Die Dateisystem-Caches und NFS arbeiten bei optimaler Datenlokalität fast ohne Netzverkehr; Durchsatz und Antwortzeit gleichen einer Standalone-Konfiguration. Mit abnehmender Datenlokalität werden zunehmend das Netzwerk und der File-Server beansprucht.

Dieser Test entspricht z.B. dem Aktivieren von Applikationen in einer normalen Client-Server Produktionsumgebung. Auch dort sind von der Festplatte Datenblöcke einzulesen. Die folgende Tabelle zeigt für einige typische Applikationen, wieviele Datenblöcke initial gelesen werden, bis die Anwendung am Bildschirm erscheint. Die beobachteten Wartezeiten spiegeln die Wechselwirkung zwischen Netzwerk, I/O-Subsystem und Virtual Memory Management wider. Bis auf *ZMAIL* verwenden alle Applikationen *Shared Libraries*, d.h. diese profitieren bei der Aktivierung auf dem NFS-Client von lokal installierten *DEC OSF/1 Libraries*.

<b>Applikation</b>	<b>Größe des Textsegments (KByte)</b>	<b>Wartezeit bei lokalem Plattenzugriff (sec)</b>	<b>Wartezeit bei NFS-Zugriff (sec)</b>	<b>Anzahl gelesener Blöcke beim Aktivieren (1 KByte Einheiten)</b>
<i>ImageMagick</i> ( <i>combine, display....</i> )	980	2	2	530
<i>ZMAIL</i>	4056	7	12	3389
<i>TGIF</i>	1274	2	3	1039
<i>EMACS</i>	2316	3	5	2059
<i>XV</i>	1964	3	3	1095

**Tabelle 1: Aktivierungszeiten einiger typischer Applikationen**

## **Testoperation 2: Suchen über NFS**

Statt des datenintensiven Lesevorgangs aus dem vorherigen Test wird beim Test "Suchen über NFS" mittels des Unix-Befehls "find" auf einem bis fünf Clients eine Operation gestartet, die auf Verarbeitung von Metadaten (hier *directory files*) ausgerichtet ist.

Dabei führen die Client-Systeme crontab-gesteuert "***find /usr/local -name 'nix' -print***" aus und durchsuchen damit gleichzeitig ein Dateisystem auf einer RZ26L (400 MByte, 16000 Inodes) des Servers. Die Systemzeiten dazu werden durch das XNTPD-Protokoll synchronisiert.

Die CPU-Belastung der beteiligten Systeme ist bei diesem Test höher als bei üblichen Datenblockoperationen; die max. Transferrate des LANs hat keinen Einfluß auf die Ausführungszeit. Das Datenvolumen ist gering, durchschnittlich sind die gesendeten RPCs und die Antworten in diesem Test nur 170 Byte lang. Limitierend wirken an dieser Stelle der durch NFS-Protokoll und den zugrundeliegenden *RPC-Transport*-Mechanismus bedingte Overhead. Eine Workstation muß zur Abarbeitung der hier gestellten Aufgabe rund 19000 RPCs absetzen.

## **Testoperation 3: Schreiben über NFS**

Beim Test "Schreiben über NFS" schreiben einer bis fünf Clients gleichzeitig über NFS (NFS V2 write) eine 10 MByte große Datei von einer lokalen RZ26 auf eine RZ58 des Servers.

Dabei führen die Client-Systeme crontab-gesteuert "***cp local\_10MB.dat nfs\_10MB.dat***" aus. Die Systemzeiten dazu werden durch das XNTPD-Protokoll synchronisiert.

Bedingt durch die NFS-Implementierung verlangen Schreiboperationen eine synchrone Abarbeitung auf dem File-Server. Die Vorteile der im Unix vorhandenen Dateisystem-Caches können an dieser Stelle nicht genutzt werden; ein File-Server (NFS-Server) bestätigt den Schreibauftrag erst, wenn die Daten physikalisch auf die Platte geschrieben sind. Nur durch zusätzliche Hardware und Systemsoftware kann ohne Abstriche bei der Datensicherheit bzw. ohne Verletzung des NFS-Protokolls bei der Abarbeitung von Schreibaufträgen eine Leistungssteigerung erzielt werden. In der Regel liegen die Transferraten beim NFS-Write deutlich unter der nach heutigem Maßstab gering einzustufenden ETHERNET-Bandbreite. Zur Leistungssteigerung wird für zahlreiche Rechnerfamilien ein NFS-Beschleuniger, *Prestoserve*, angeboten. Dieser NFS-Beschleuniger optimiert die Abarbeitung von NFS-Write und Meta-Operationen, im Prinzip alle synchronen Dateisystemoperationen, und ist sogar bei einer ETHERNET-basierten Verbindung eine wirkungsvolle Erweiterung zur Durchsatzsteigerung.

In dem ALPHA AXP System DEC 3000-M400 ist *Prestoserve* installiert und bei allen Tests aktiviert. In der Betriebssystemversion DEC OSF/1 V3.0 ist zur netzwerkweiten Unterstützung von 64-Bit-Dateisystemen die NFS-Version 3.0 implementiert; diese neue NFS-Version beinhaltet u.a. Mechanismen zur Durchsatzsteigerung bei Schreiboperationen, die mittelfristig das Einsatzspektrum von *Prestoserve* voraussichtlich reduzieren werden. Derzeit existieren noch Schwierigkeiten bei FDDI-Verbindungen und der Interoperabilität mit dem NFS-Automounter.

## Testergebnisse

### 1. Gemeinsames Ethernet für die Workstation-Gruppe und ca. 200 weitere Netzwerkteilnehmer

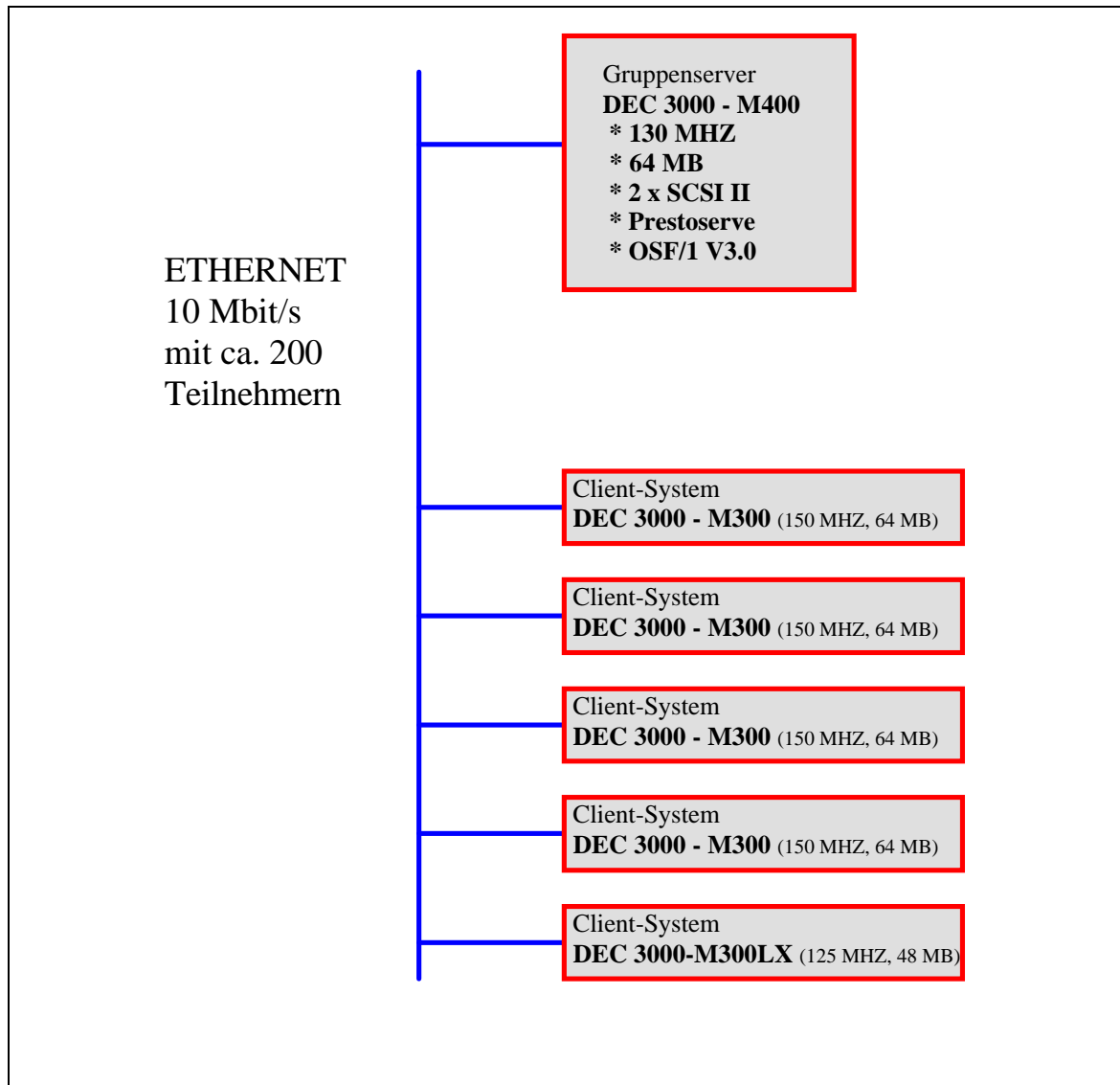


Abbildung 2: Netzwerkkonfiguration 1

In dieser Netzwerkkonfiguration teilt sich die Workstation-Gruppe ein Ethernet mit einer großen Anzahl weiterer Netzwerkteilnehmer (Workstations, PCs und X-Terminals). Diese Konfiguration ist als die ungünstigste für die Workstation-Gruppe anzusehen, da um die Ethernetbandbreite von 10 Mbit/s mit einer großen Anzahl weiterer Maschinen konkurriert werden muß.



### 1. a. Leseoperation bei einem gemeinsamen Ethernet für die Workstation-Gruppe und weitere Netzwerkteilnehmer

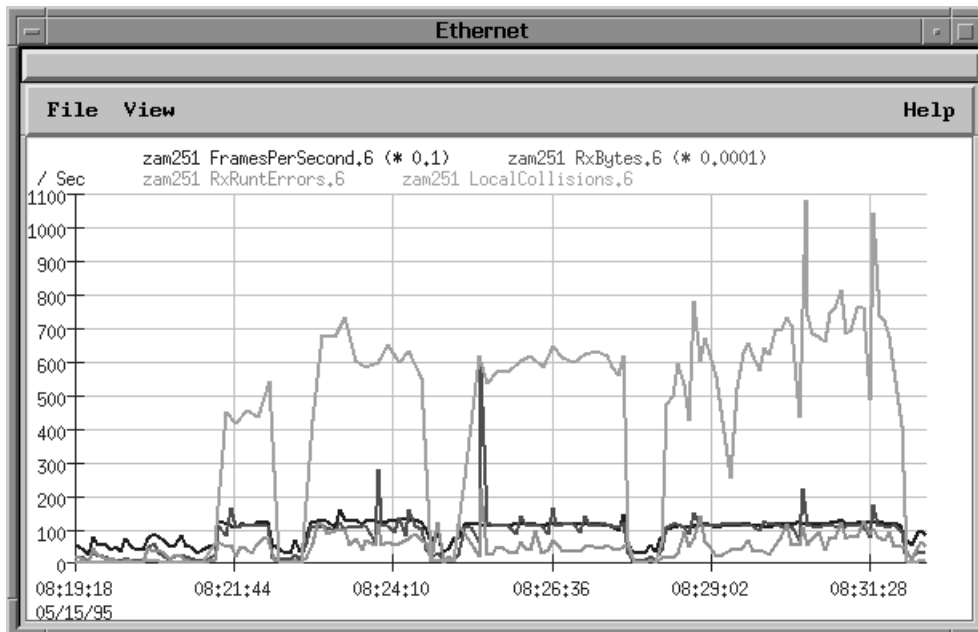


Abbildung 3: Ethernet-Statistik während des Tests 1. a.

Die Leseoperation wurde in dieser Konfiguration nur mit einem bis vier Clients gleichzeitig durchgeführt, um die Einschränkungen für die anderen Netzwerkteilnehmer in diesem Produktionsnetz in Grenzen zu halten (während der Tests war tatsächlich für andere Teilnehmer ein Arbeiten über das Ethernet so gut wie unmöglich).

### 1. b. Suchoperation bei einem gemeinsamen Ethernet für die Workstation-Gruppe und weitere Netzwerkteilnehmer

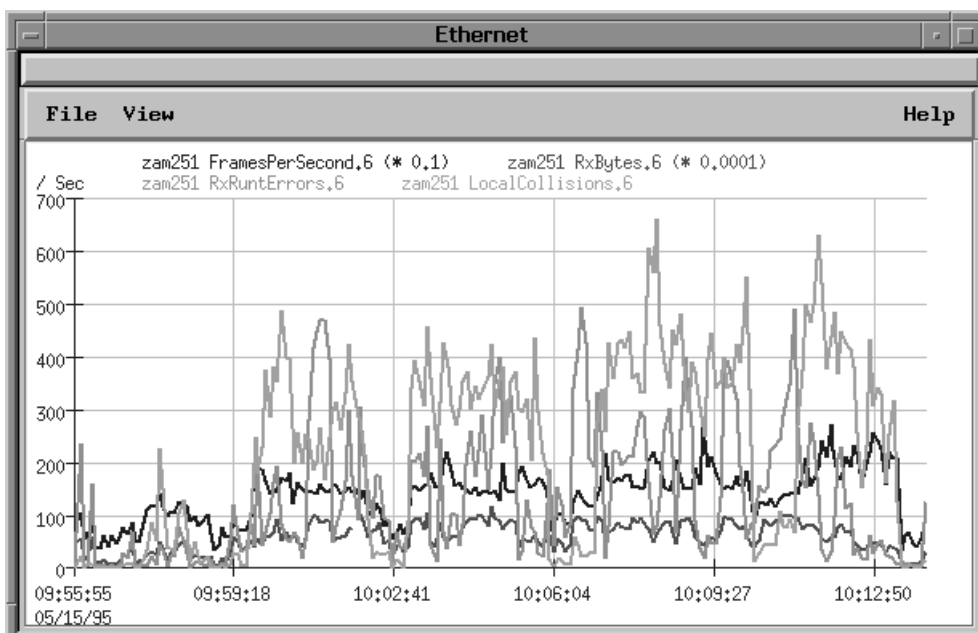


Abbildung 4: Ethernet-Statistik während des Tests 1. b.

### 1. c. Schreiboperation bei einem gemeinsamen Ethernet für die Workstation-Gruppe und weitere Netzwerkteilnehmer

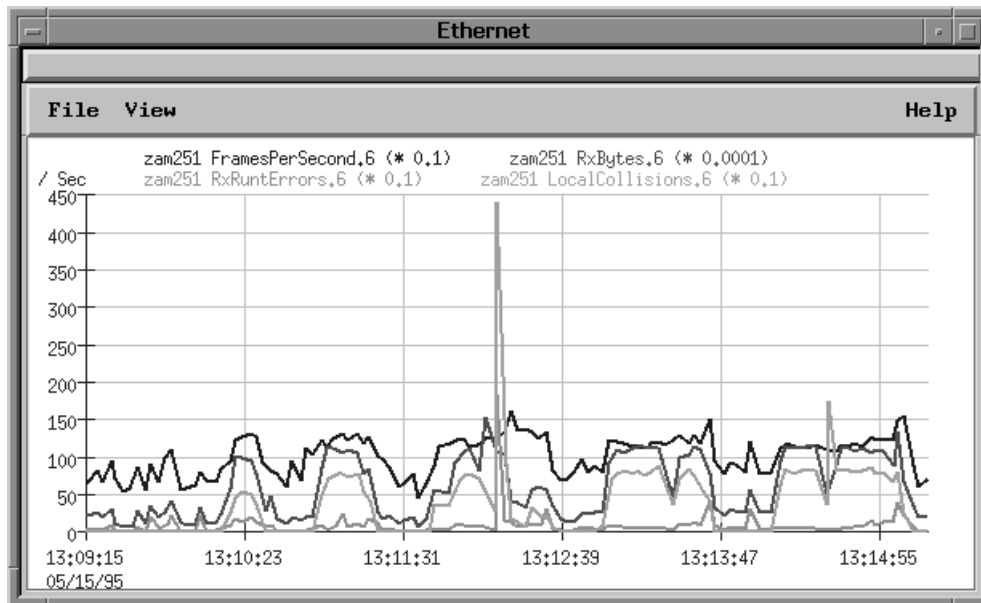


Abbildung 5: Ethernet-Statistik während des Tests 1. c.

### 1. d. Zusammenfassung der Ergebnisse zur Netzwerkkonfiguration 1

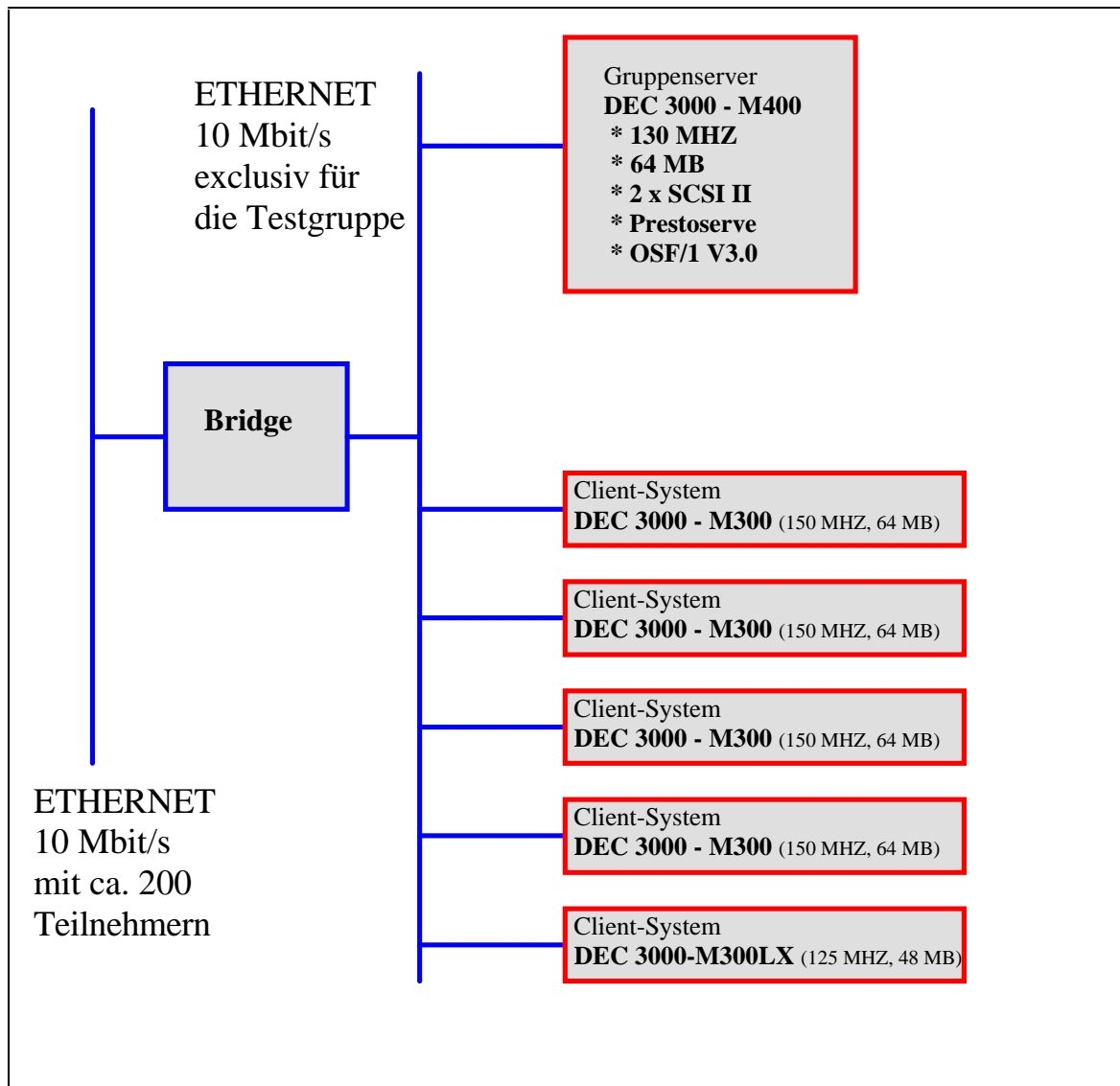
Bei diesem Test müssen sich die Testmaschinen das Ethernet mit weiteren ca. 200 Teilnehmern teilen. Entsprechend hoch ist die Anzahl der Kollisionen ( $> 1000/\text{sec}$ ), insbesondere während der Leseoperation mit ihrem kontinuierlichen Datenstrom vom Server zum Client.

Die Auslastung des Ethernet während der Leseoperation liegt bei annähernd 90%, weist aber Schwankungen aufgrund der Konkurrenzsituation auf dem Netz auf.

Die Ausführungszeit der Leseoperation ist um ca. 10-15% länger als die vergleichbarer Operationen mit der Netzwerkkonfiguration 2 (separates Ethernet für die Workstation-Gruppe).

Während der Tests dominieren die Testmaschinen das Ethernet, so daß ein Arbeiten mit dem Ethernet für die anderen Netzwerkteilnehmer kaum möglich ist.

## 2. Separates Ethernet für die Workstation-Gruppe



**Abbildung 6: Netzwerkkonfiguration 2**

In dieser Netzwerkkonfiguration verfügt die Workstation-Gruppe über ein eigenes Ethernet mit einer Bandbreite von 10 Mbit/s. Als trennende Komponente zu einem weiteren Ethernet mit „Produktionsbetrieb“ und einer größeren Anzahl von Teilnehmern (ca. 200) wird eine Bridge eingesetzt. Diese transferiert nur Ethernetframes zwischen den Netzen, die als Absender- oder Zieladresse eine der Workstations der Testgruppe hat. Insbesondere isoliert sie die Testgruppe von dem nicht relevanten Verkehr zwischen den anderen Netzwerkteilnehmern im „Produktionsnetz“.

## 2. a. Leseoperation bei einem separaten Ethernet für die Workstation-Gruppe

Anzahl der Client-Systeme	NFS Transferrate pro Client, READ (KByte/s)	IDLE Time Gruppenserver (%)	Disk Transferrate Gruppenserver (KByte/s)	NFS Slow Server (badxid)
1	1000	85	1000	0
2	575	85	1020	0
3	393	82	1132	0
4	284	83	1122	0
5	226	75	1112	0
Server	-	85	3270	-

Tabelle 2: Performance-Werte auf den Workstations zum Test 2. a.

Anzahl Client Systeme + (NFS rpcs)	ETHERNET Auslastung (%)	ETHERNET Output Errors (Excessive Collisions)	ETHERNET Collisions
1 (6140 rpcs)	78	0	2
2 (12280 rpcs)	89	5	5000
3 (18420 rpcs)	92	93	12808
4 (24560 rpcs)	88	170	27115
5 (30700 rpcs)	88	318	41944

Tabelle 3: Ethernet-Statistik des Gruppenservers zum Test 2. a.

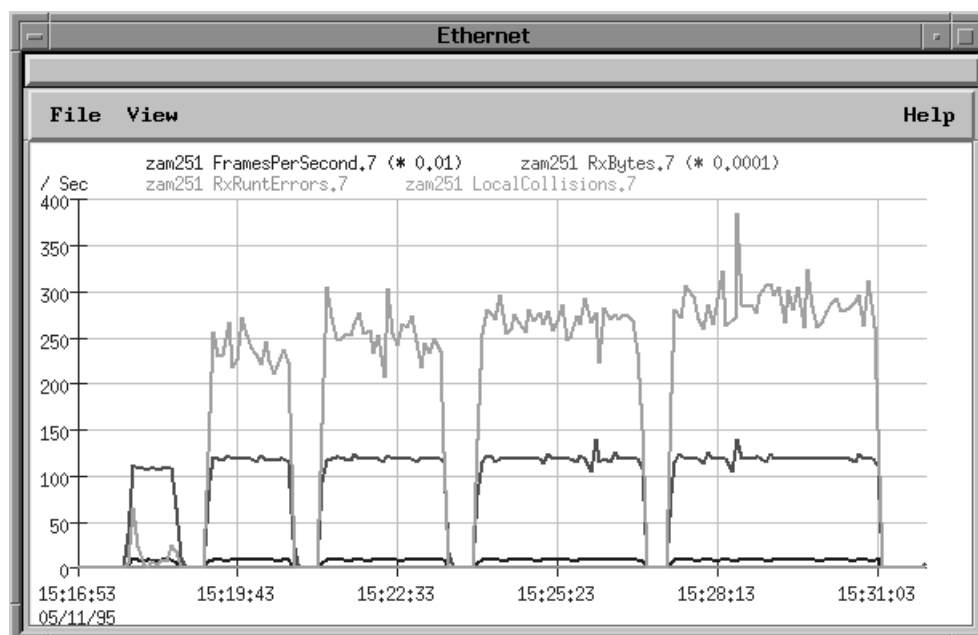


Abbildung 7: Ethernetstatistik des Tests 2. a.

## 2. b. Suchoperation bei einem separaten Ethernet für die Workstation-Gruppe

Anzahl der Client-Systeme	find-Ausführungszeit pro Client (sec)	IDLE Time Gruppenserver (%)	Disk Transferrate Gruppenserver (KByte/s)	NFS Slow Server (badxid)
1	67	78	103	0
2	67	59	103	0
3	70	36	100	0
4	76	17	91	0
5	84	12	88	0
Server	21	71	244	-

Tabelle 4: Performance-Werte auf den Workstations zum Test 2. b.

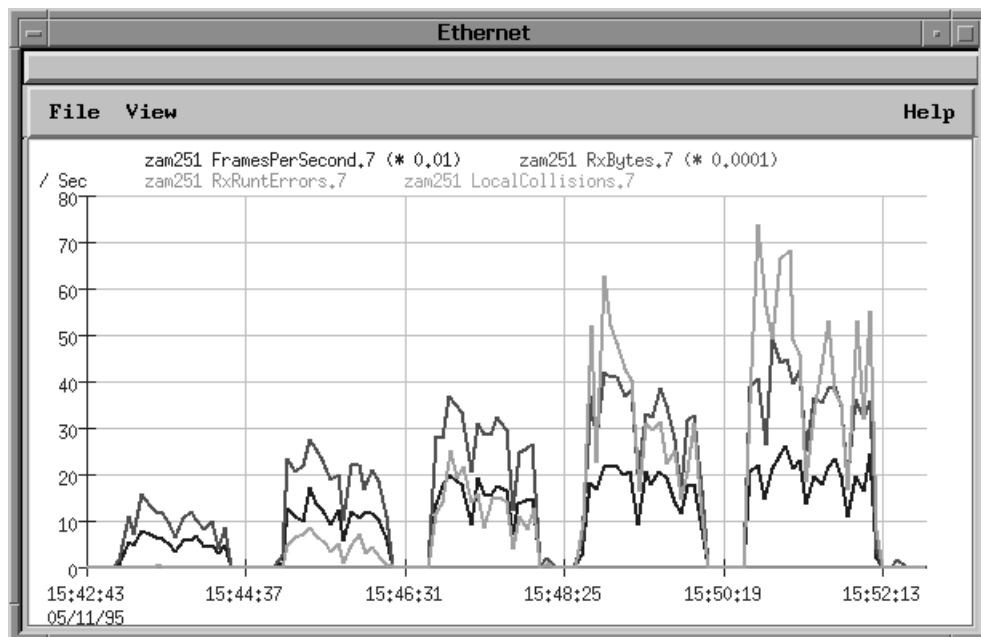


Abbildung 8: Ethernetstatistik des Tests 2. b.

## 2. c. Schreiboperation bei einem separaten Ethernet für die Workstation-Gruppe

Anzahl der Client-Systeme	NFS Transferrate pro Client, WRITE (KByte/s)	IDLE Time Gruppenserver (%)	Disk Transferrate Gruppenserver (KByte/s)	NFS Slow Server (badxid)
1	932	61	935	0
2	563	66	1078	0
3	383	58	1129	0
4	294	61	1085	0
5	227	50	1075	0
Server	-	63	2100	-

Tabelle 5: Performance-Werte auf den Workstations zum Test 2. c.

Anzahl Client Systeme + (NFS rpcs)	ETHERNET Auslastung (%)	ETHERNET Output Errors (Excessive Collisions)	ETHERNET Collisions
1 (1230 rpcs)	72	0	10
2 (2460 rpcs)	87	0	513
3 (3690 rpcs)	89	13	1670
4 (4920 rpcs)	91	58	5752
5 (6150 rpcs)	88	156	9580

Tabelle 6: Ethernet-Statistik des Gruppenservers zum Test 2. c.

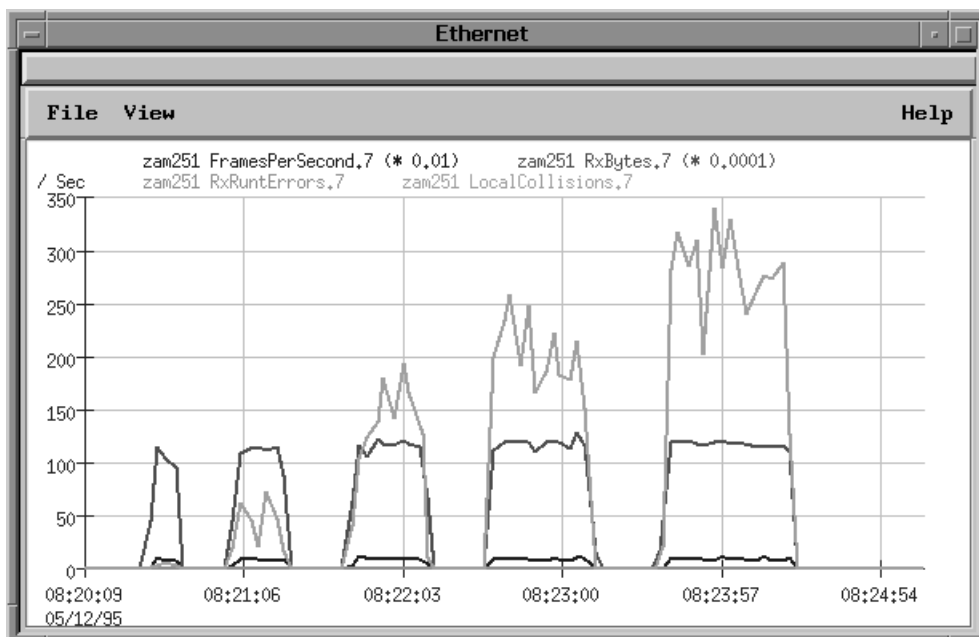


Abbildung 9: Ethernet-Statistik des Tests 2. c.

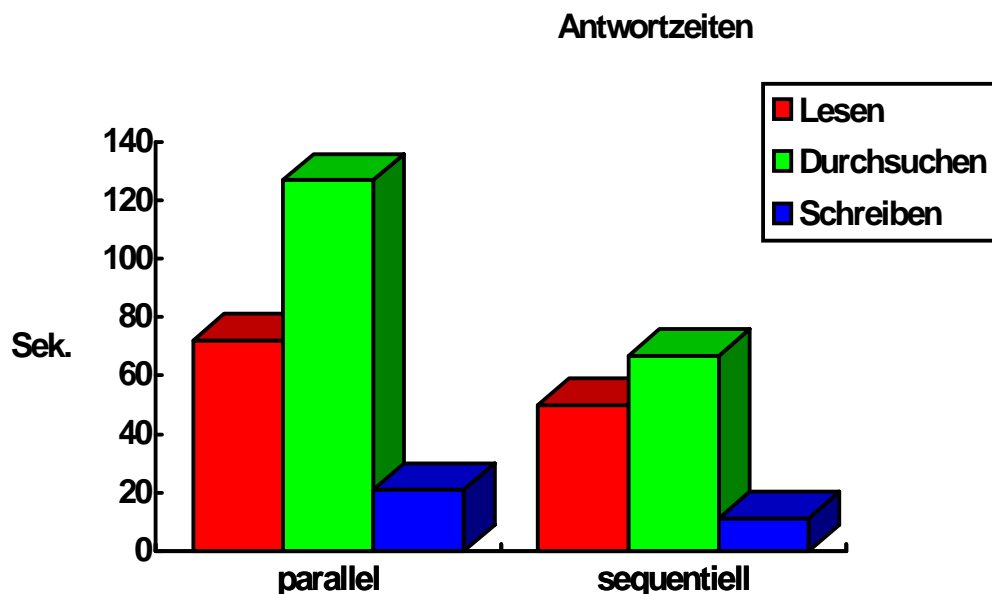
## 2. d. Zusammenfassung der Ergebnisse zur Netzwerkkonfiguration 2

Bei dieser Netzwerkkonfiguration liegt die Auslastung des Ethernet beim Lesetest kontinuierlich bei ca. 95%. Da diese Auslastung bereits beim Lesevorgang eines Clients erreicht wird, nimmt die Ausführungszeit der Tests mit 2 bis 5 Clients linear zu. Die Bandbreite des Ethernet ist in diesem Fall die den Datendurchsatz begrenzende Komponente.

Auch bei diesen Tests treten, allerdings erst bei Operationen mit mehr als einem Client, vermehrt Kollisionen auf. Die Anzahl der Kollisionen ist allerdings nicht so hoch wie bei der Netzwerkkonfiguration 1. Das aus Testgruppensicht hinter der Bridge liegende Produktionsethernet wird, im Gegensatz zur Netzwerkkonfiguration 1, durch die Testoperationen nicht beeinflusst.

## 2. e. Operationen parallel

Abschließend werden zu dieser Testumgebung die Ergebnisse bei der gleichzeitigen Ausführung von *NFS-Lesen*, *NFS-Schreiben* und *NFS-Durchsuchen* dargestellt und mit der aufeinanderfolgenden Ausführung der gleichen Operationen auf einem NFS-Client verglichen. Dabei wurden parallel zum Test *Durchsuchen* eine 50 MByte Datei gelesen und in kurzen Abständen dreimal eine 10 MByte Datei geschrieben.



**Abbildung 10: Antwortzeiten bei paralleler und sequentieller Ausführung der Tests**

Die ohnehin hohe Antwortzeit bei der Aufgabe *Durchsuchen* wird hier durch die *Lesen*- und *Schreiben*-bedingte ETHERNET-Last drastisch verschlechtert. Eine Reduzierung der Antwortzeiten um rund 30% bei der Aufgabe *Durchsuchen* zeigen derzeit erste Tests mit NFS V3; neu implementierte RPCs wie z.B. *readdir+* reduzieren die Anzahl der benötigten RPC-Aufrufe von über 19000 auf rund 3500.

An dieser Stelle werden die Antwortzeiten aus den bisherigen Tests *Lesen/Durchsuchen/Schreiben* für die entsprechende Anzahl gleichzeitig aktiver NFS-Clients zusammengefaßt, um einen direkten Vergleich mit dem oben beschriebenen Szenario zu ermöglichen.

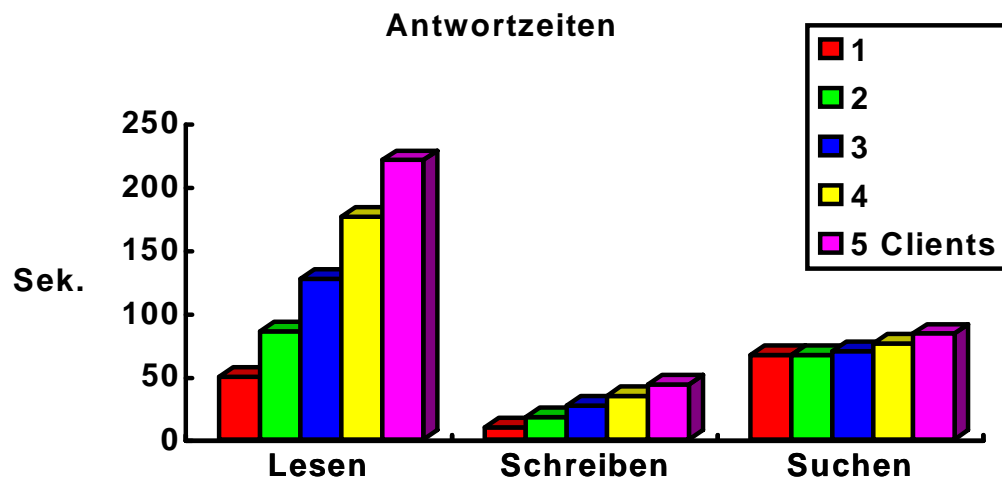


Abbildung 11: Antwortzeiten aller Tests mit der Netzwerkkonfiguration 2.



### 3. Einsatz eines Ethernet-FDDI-Switches zwischen Gruppenserver und Clients

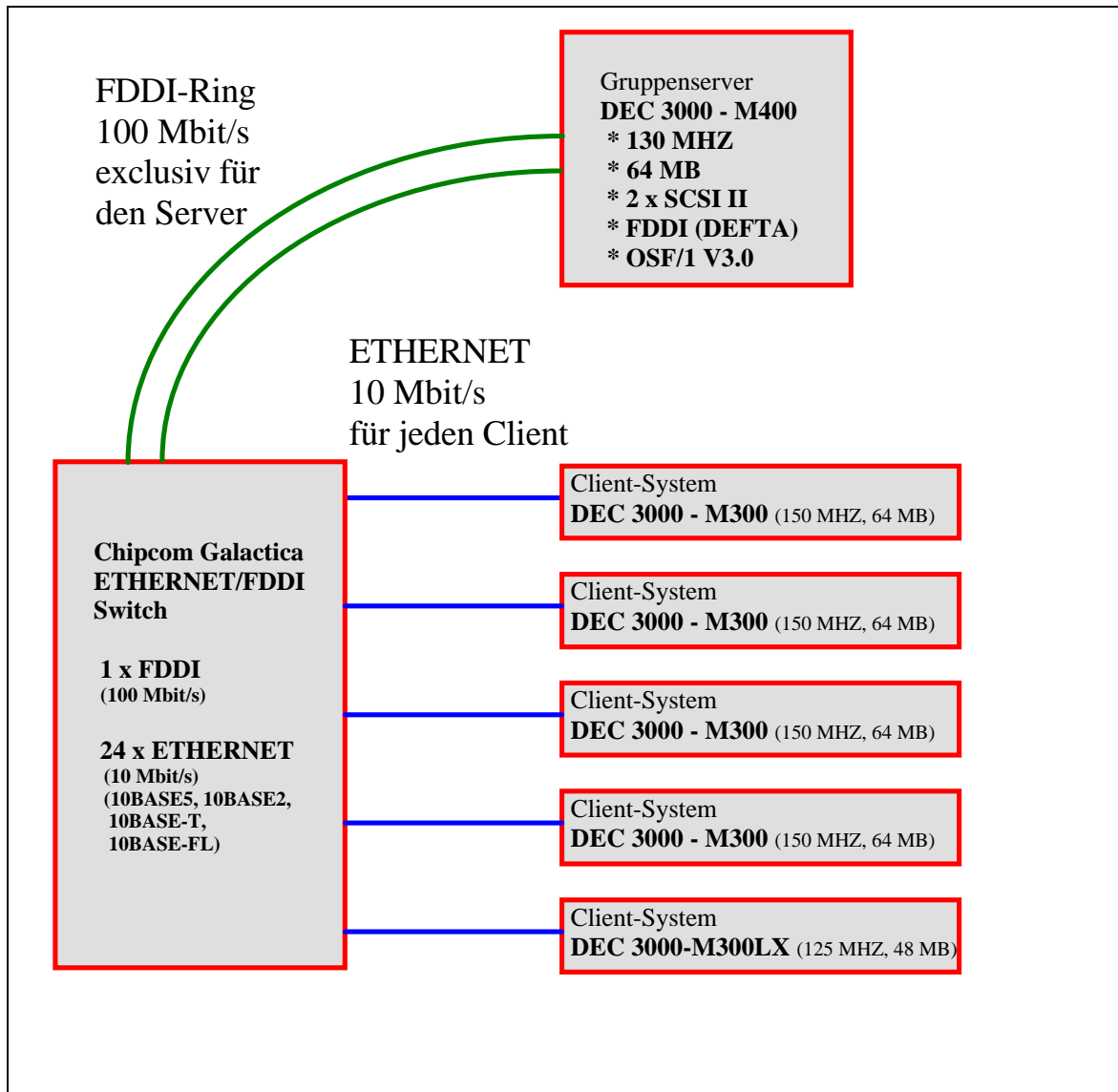


Abbildung 12: Netzwerkkonfiguration 3

In dieser Netzwerkkonfiguration wird zwischen den Clients und dem Server ein Ethernet-FDDI-Switch eingesetzt. Dieser ermöglicht jedem Client, über einen Ethernet-Switch-Port mit jeweils 10 MBit/s zu kommunizieren.

Der Anschluß des Servers erfolgt über einen FDDI-Port des Switches mit exklusiven 100 MBit/s.

Diese Konfiguration ermöglicht den gleichzeitigen Zugriff der Clients auf den Server mit  $n \times 10$  MBit/s (aufgrund von CPU- und netzwerkadapterbedingten Einschränkungen auf der Serverseite ist  $n < 6$  anzunehmen, wie die folgenden Tests zeigen werden (insbesondere Test 3. d.)).

### 3. a. Schreiboperation beim Einsatz eines Ethernet-FDDI-Switches

Anzahl der Client-Systeme	NFS Transferrate pro Client, READ (KBytes/s)	IDLE Time Gruppenserver (%)	Disk Transferrate Gruppenserver (KByte/s)	NFS Slow Server (badxid)
1	950	85	950	-
2	948	70	1888	-
3	931	60	2675	-
4	736	55	2793	-
5	602	56	2709	-
Server	-	85	3270	-

Tabelle 7: Performance-Werte auf den Workstations zum Test 3. a.

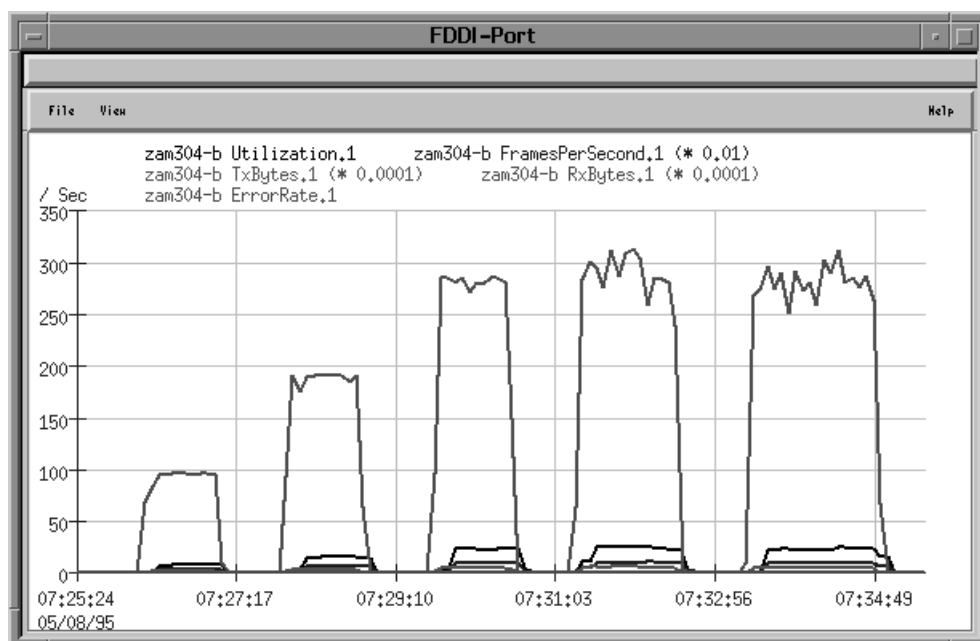


Abbildung 13: FDDI-Statistik zum Test 3. a.

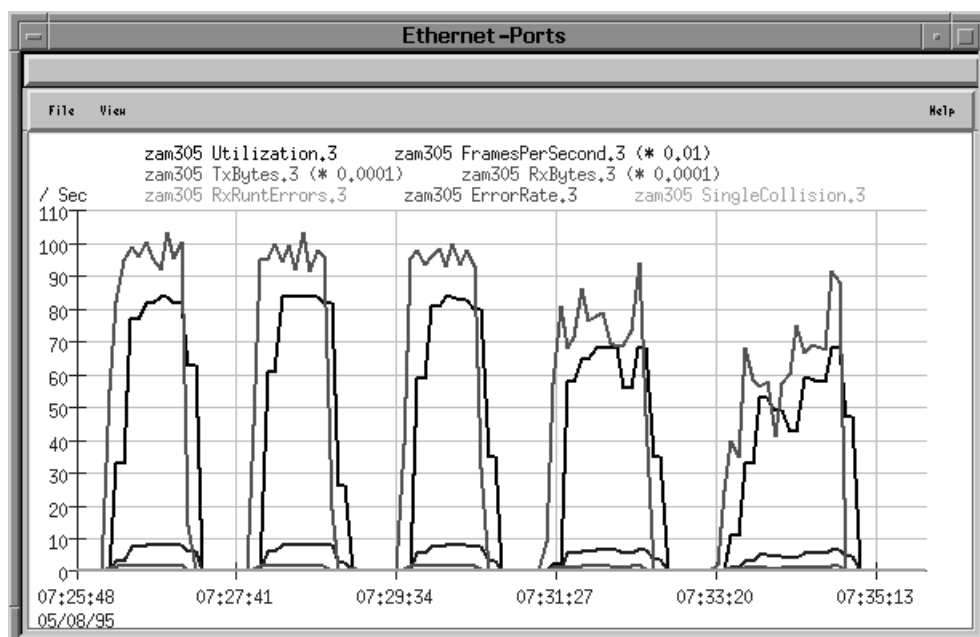


Abbildung 14: Ethernet-Statistik des ersten Switchports zum Test 3. a.

### 3. b. Suchoperation beim Einsatz eines Ethernet-FDDI-Switches

Anzahl der Client-Systeme	<i>find</i> -Ausführungszeit pro Client (sec)	<i>IDLE Time</i> Gruppenserver (%)	Disk Transferrate Gruppenserver (KByte/s)	NFS <i>Slow Server</i> (badxid)
1	78	80	101	-
2	80	54	92	-
3	79	30	92	-
4	81	15	87	-
5	87	11	78	-
Server	21	71	244	-

Tabelle 8: Performance-Werte auf den Workstations zum Test 3. b.

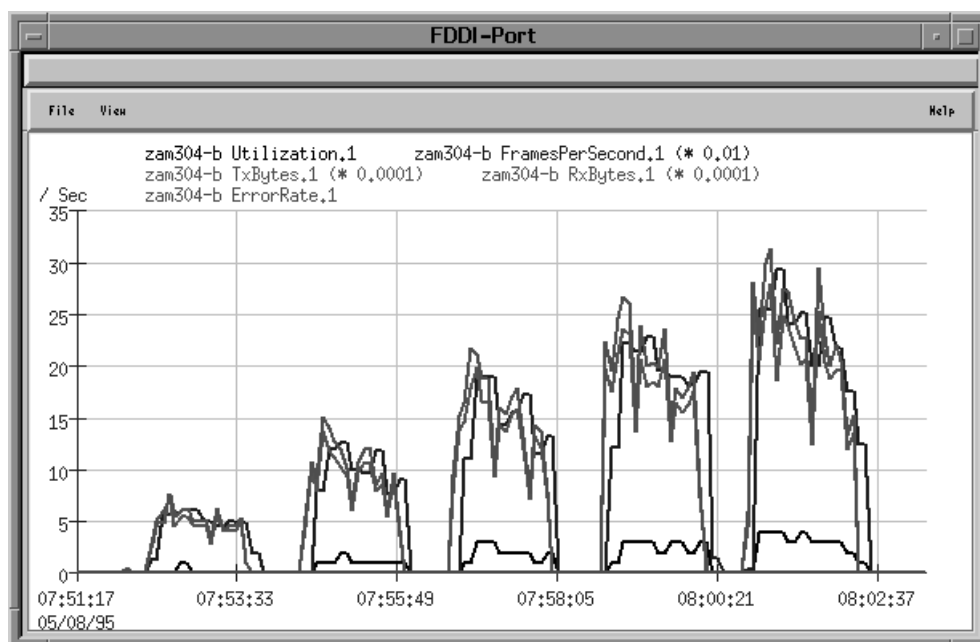


Abbildung 15: FDDI-Statistik zum Test 3. b.

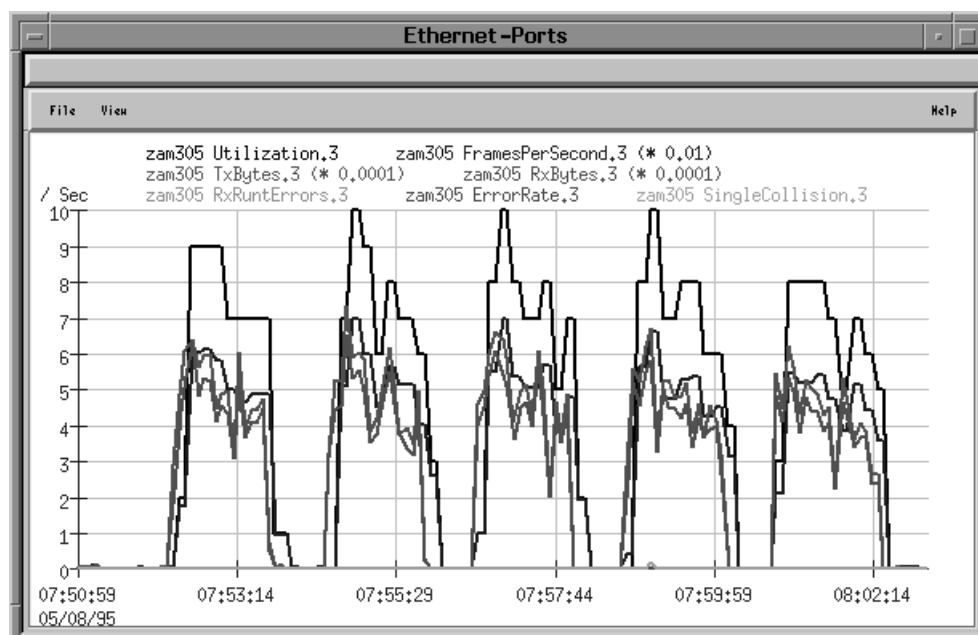


Abbildung 16: Ethernet-Statistik des ersten Switchports zum Test 3. b.

### 3. c. Schreiboperation beim Einsatz eines Ethernet-FDDI-Switches

Anzahl der Client-Systeme	NFS Transferrate pro Client, WRITE (KByte/s)	IDLE Time Gruppenserver (%)	Disk Transferrate Gruppenserver (KByte/s)	NFS Slow Server (badxid)
1	1013	60	1000	-
2	930	40	1805	-
3	640	37	1850	-
4	471	42	1694	-
5	365	46	1580	-
Server	-	63	2100	-

Tabelle 9: Performance-Werte auf den Workstations zum Test 3. c.

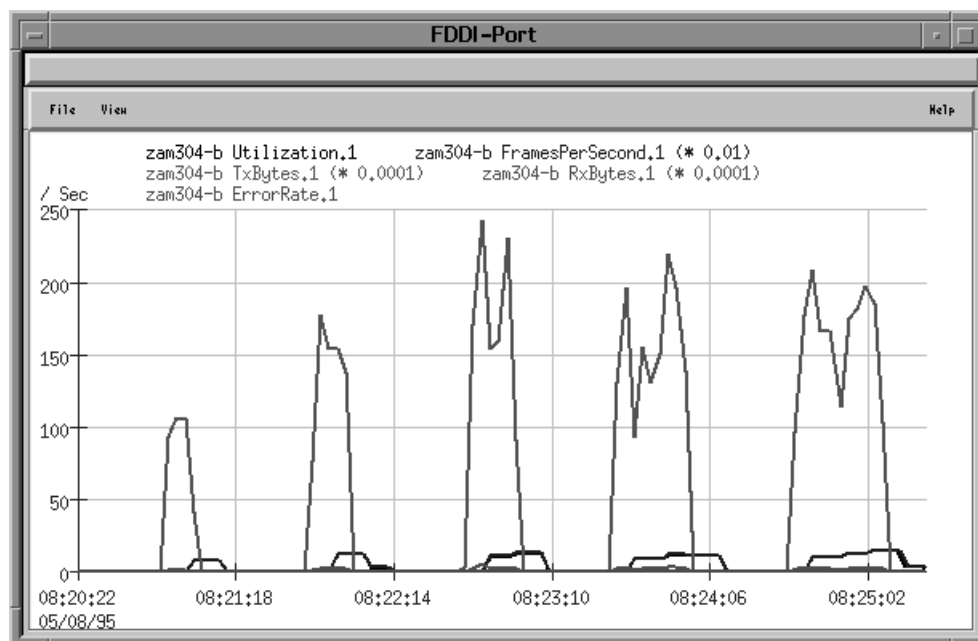


Abbildung 17: FDDI-Statistik zum Test 3. c.

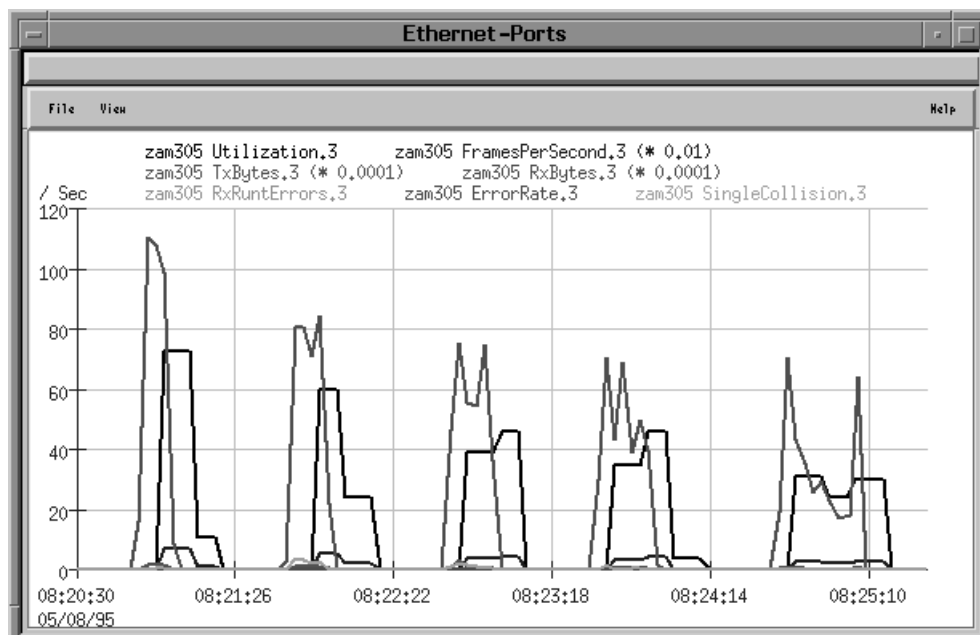


Abbildung 18: Ethernet-Statistik des ersten Switchports zum Test 3. c.

### 3. d. TCPSPRAY beim Einsatz eines Ethernet-FDDI-Switches

Anzahl der Client-Systeme	Transferrate pro Client in KByte/s
1	1090
2	1070
3	1053
4	1006
5	872

Tabelle 10: Performance-Werte auf den Workstations zum Test 3. d.

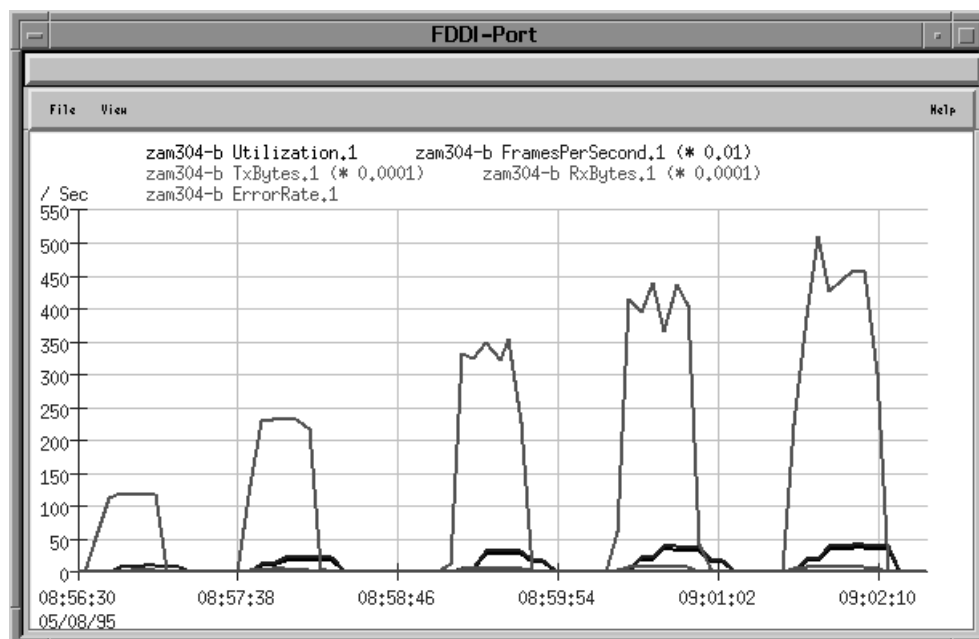


Abbildung 19: FDDI-Statistik zum Test 3. d.

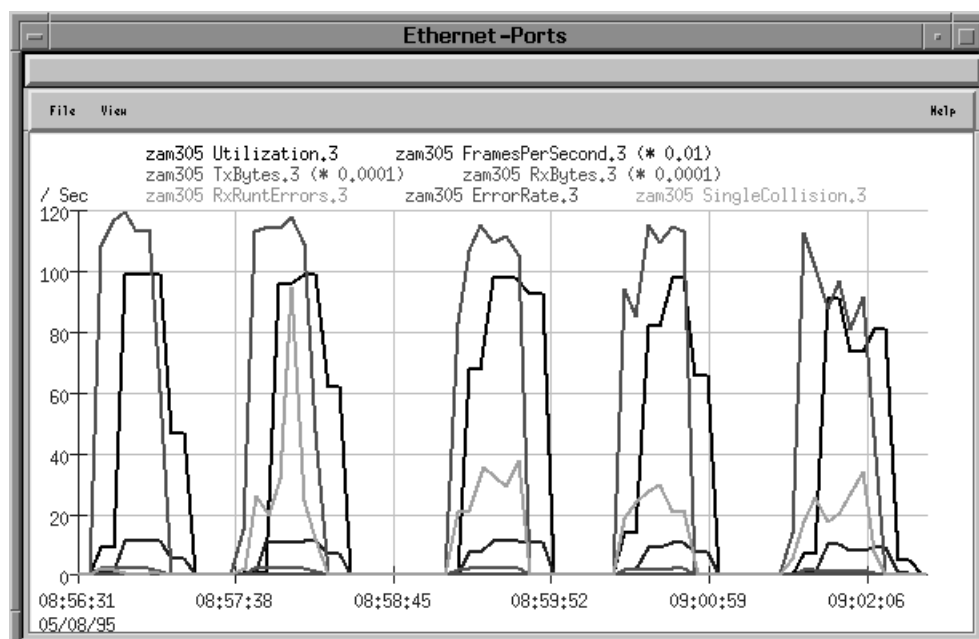


Abbildung 20: Ethernet-Statistik des ersten Switchports zum Test 3. d.

### 3. e. Zusammenfassung der Ergebnisse zur Netzwerkkonfiguration 3

Bei dieser Netzwerkkonfiguration liegt die Anfangsauslastung des ersten Ethernetsegmentes beim Lese-  
test mit einem Client mit ca. 85% etwas niedriger als beim Test 2. Dies läßt sich durch den zwischen  
dem Client und dem Server liegenden Switch erklären, der jeden Ethernetframe filtern und aktiv weiter-  
leiten muß.

Im weiteren Verlauf des Tests mit bis zu fünf Clientlesevorgängen gleichzeitig wird dann der große  
Vorteil dieser Konfiguration sichtbar: bis zu einer Anzahl von 3 Clients wächst der Datendurchsatz des  
Systems linear auf annähernd 30 MBit/s an; erst ab 4 Clients gleichzeitig tritt eine Konkurrenzsituation  
auf, und die Ausführungszeit der Gesamtoperation verlängert sich leicht. Dies ist auf die Plattenperi-  
pherie des Servers zurückzuführen, wie ein Vergleich mit dem Test 3. d. zeigt, bei dem ein Memory to  
Memory Transfer zwischen Server und Clients erfolgte, bei dem der Datendurchsatz sich bis annähernd  
50 MBit/s steigern ließ.

An dieser Stelle seien als Zusammenfassung nochmals die Antwortzeiten aus den bisherigen Tests *Le-  
sen/Durchsuchen/Schreiben/TCPSPRAY* für die entsprechende Anzahl gleichzeitig aktiver NFS-Clients  
dargestellt:

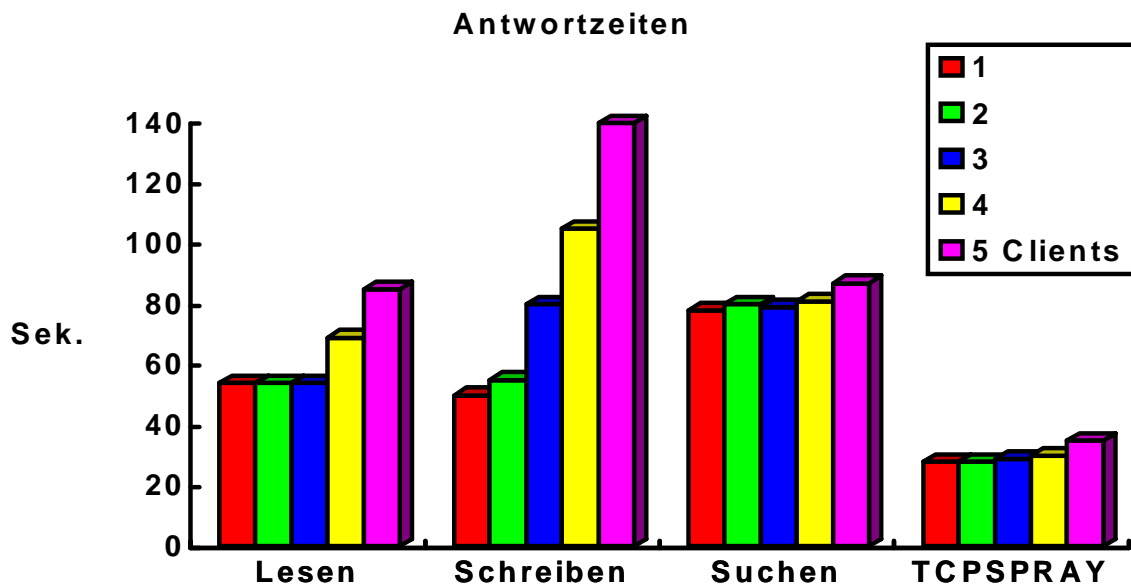


Abbildung 21: Antwortzeiten aller Tests mit der Netzwerkkonfiguration 3

#### 4. Anbindung eines Clients über FDDI an den Server

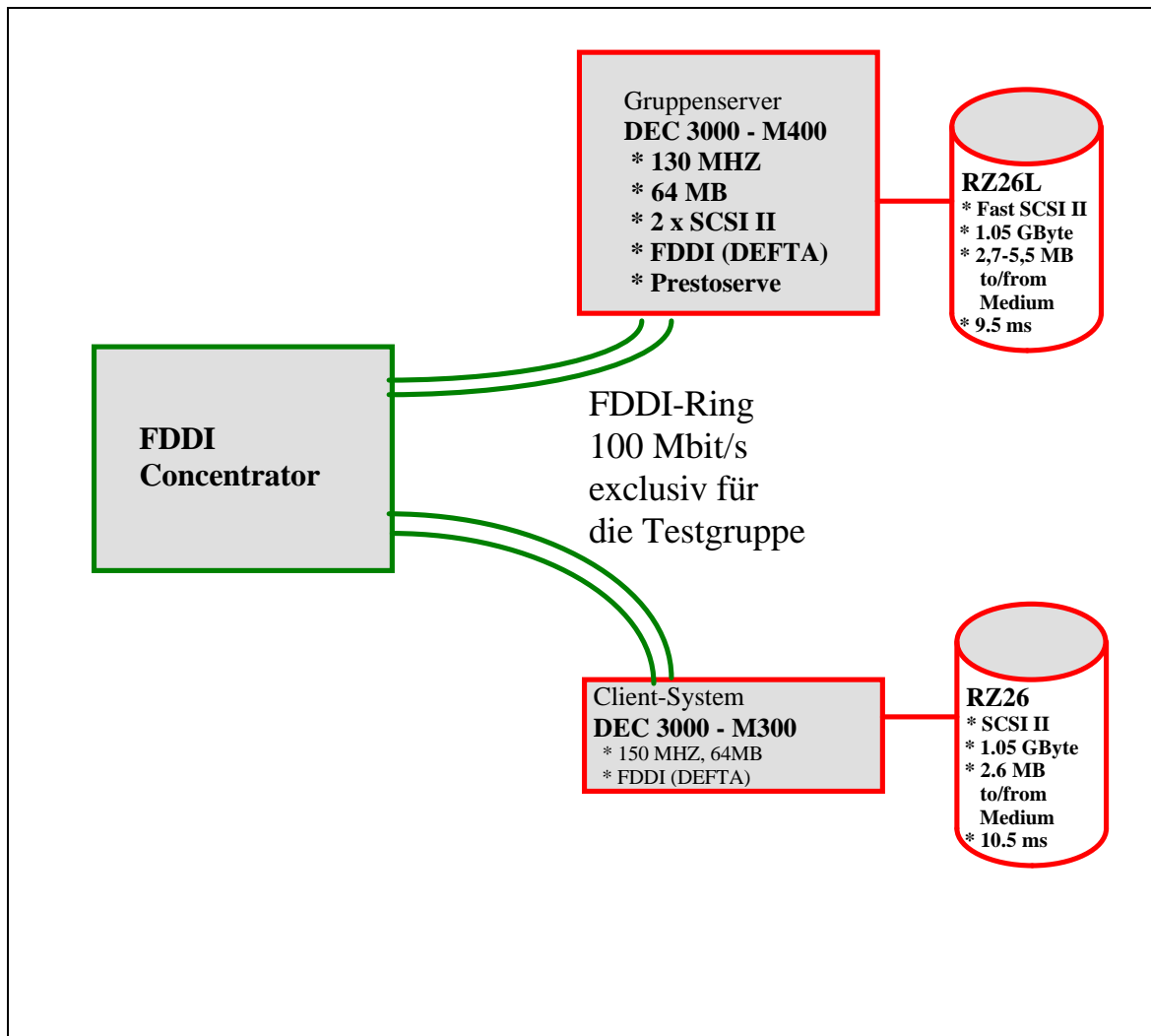


Abbildung 22: Netzwerkkonfiguration 4

Bei diesem Test wurde die "Idealkonfiguration" für eine Workstation-Gruppe gewählt:

Server und Client sind an einem exclusive FDDI-Ring angeschlossen.

Aus Kostengründen konnten die Tests zu dieser Konfiguration nur mit einem Client gefahren werden.

##### 4. a. Performance-Werte beim Einsatz eines FDDI-Ringes zwischen Client und Server

Applikationen	Transferrate (KByte/s)	IDLE Time Gruppenserver (%)
NFS READ (disk to memory)	2553	65
NFS READ (disk to disk)	1890	75
NFS WRITE (disk to disk)	1536	69
TCPSPRAY (blocksize=8K)	6661	50
FTP (disk to memory)	2800	-
FTP (disk to disk)	2400	-

Tabelle 11: Performance-Werte auf den Workstations zum Test 4

#### **4. b. Zusammenfassung der Ergebnisse zur Netzwerkkonfiguration 4**

Die Transferraten zu diesen Tests erreichen bis zu 53 MBit/s bei einem Memory to Memory Transfer mit 8 KByte Blockung und spiegeln die Leistungsfähigkeit dieser Netzwerkkonfiguration insbesondere bei Spitzenanforderungen eines einzelnen Clients wider, die bei einer Switchlösung durch den Ethernetanschluß des Clients auf 10 MBit/s begrenzt wird.

Bezüglich des Gesamtdurchsatzes des Systems bei Anforderungen durch mehrere Clients unterscheidet sich diese Konfiguration nicht wesentlich von der Switchlösung (Netzwerkkonfiguration 3).



## Preisvergleich

Die verschiedenen in den Tests vorgestellten Szenarios zur Vernetzung von Workstation-Gruppen unterscheiden sich nicht nur deutlich in ihrer Leistungsfähigkeit sondern auch in den mit der Realisierung verbundenen Kosten.

Im folgenden werden die Kosten für die jeweilige Hardware (sowohl auf Workstationseite als auch für aktive Netzwerkkomponenten (nicht incl. Kabelinfrastruktur)) abgeschätzt. Die tatsächlichen Kosten können gegenüber den Abschätzungen durch die Auswahl spezieller Netzwerkkomponenten (z.B. mit erweiterter Funktionalität oder erweiterten Managementmöglichkeiten) oder spezieller Hersteller durchaus deutlich abweichen.

Bei den FDDI-Komponenten wird von einer Verkabelung auf Glasfaserbasis ausgegangen. Bei Einsatz von UTP- oder S-UTP-Kabeln sind die Kosten der FDDI-Komponenten ca. 30% niedriger als angegeben.

<i>Konfiguration</i>	<i>Preis der Kommunikationskomponente pro Workstation</i>	<i>Preis der Netzwerkkomponenten pro Workstation</i>	<i>Preis der kompletten Ausstattung einer Workstation-Gruppe mit einem Server und 8 Clients</i>
<b>1: gemeinsames Ethernet mit 10Base2-Anschlüssen (Thinwire)</b>	0,- DM bis 300,- DM	20,- DM	180,- DM bis 2.880,- DM
<b>2: separates Ethernet für die Workstation-Gruppe mit 10BaseT-Anschlüssen (Einsatz einer Bridge)</b>	0,- DM bis 80,- DM	400,- DM	3.600,- DM bis 4.320,- DM
<b>3: Einsatz eines Ethernet-FDDI-Switches zwischen Server und Clients (Clients mit 10BaseT-Anschlüssen)</b>	Client: 0,- DM bis 80,- DM Server: 2.500,- DM	2.300,- DM	23.840,- DM
<b>4: separater FDDI-Ring für Server und Clients</b>	2.500,- DM	2.260,- DM	42.840,- DM

**Tabelle 12: Kosten der vorgestellten Netzwerkkonfigurationen**

## Hintergrundinformationen zu NFS

Zwischen den in den Tabellen zu den Tests gezeigten Durchsatzraten und der Antwortzeit besteht hier im NFS-Kontext folgender einfacher funktionaler Zusammenhang:

$$\text{Durchschnittliche\_RPC\_Antwortzeit} = \text{NFS\_RSIZE} / \text{Durchschnittliche\_Transferrate}$$

Der Parameter *NFS\_RSIZE* legt die Blocklänge für die Leseanforderungen fest (im Standardfall 8 KByte).

Die Qualität der Netzwerkverbindung hat erheblichen Einfluß auf die NFS-Performance. Der NFS-Client erwartet zu jedem RPC (NFS-Remote Procedure Call), der UDP (User Datagram Protocol) basierend arbeitet, eine Bestätigung/Antwort durch den NFS-Server. Diese Bestätigung kann ausbleiben, wenn beispielsweise Pakete auf dem Netz verloren gehen oder eine fehlerhafte Prüfsumme (ETHERNET CRC oder UDP Checksum) vorliegt. Der NFS-Client muß in diesem Fall nach einer bestimmten Wartezeit (*timeout*) eine erneute Übertragung (*retransmission*) starten. In der folgenden Beispielrechnung wird aufgezeigt, wie hoch die max. tolerierbare NFS-Retransmission-Rate bei 20% Performance-Verlust liegen darf. Die durchschnittliche Antwortzeit auf einen RPC wird auf 10 Millisekunden (Erfahrungswert) angesetzt :

**n** == Anzahl der benötigten RPCs

**T<sub>timeout</sub>** == NFS-Timeout-Parameter, festgelegt im Mount-Protokoll auf 1.1 Sekunden

**T<sub>rpc</sub>** == mittlere Antwortzeit pro RPC, geschätzt auf 10 Millisekunden

**t<sub>rpc</sub>** == Zunahme der Antwortzeit 20%, d.h. 2 Millisekunden

**R<sub>retrans</sub>** == Anzahl der NFS-Retransmissions

$$n * ( T_{\text{rpc}} + t_{\text{rpc}} ) = n * T_{\text{rpc}} + R_{\text{retrans}} * T_{\text{timeout}}$$

⇔

$$n * t_{\text{rpc}} = R_{\text{retrans}} * T_{\text{timeout}}$$

⇔

$$t_{\text{rpc}} / T_{\text{timeout}} = R_{\text{retrans}} / n \quad (\text{Retransmission-Quote})$$

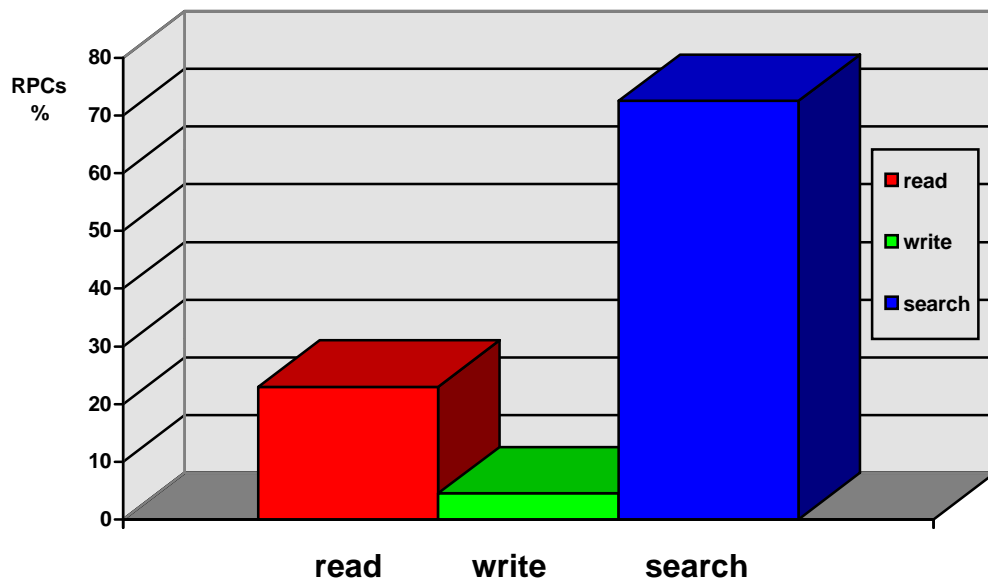
$$\Rightarrow \text{max. } 0.18 \% \text{ Retransmissions sind zulässig}$$

Pro NFS-RPC sind bei *rsize=8KByte* auf *ETHERNET*-Ebene 7 Pakete zu übertragen. Falls eines dieser Pakete fehlerhaft ist, muß der NFS-Client den RPC erneut senden; damit ergibt sich als Schranke für den *ETHERNET*-Fehler:

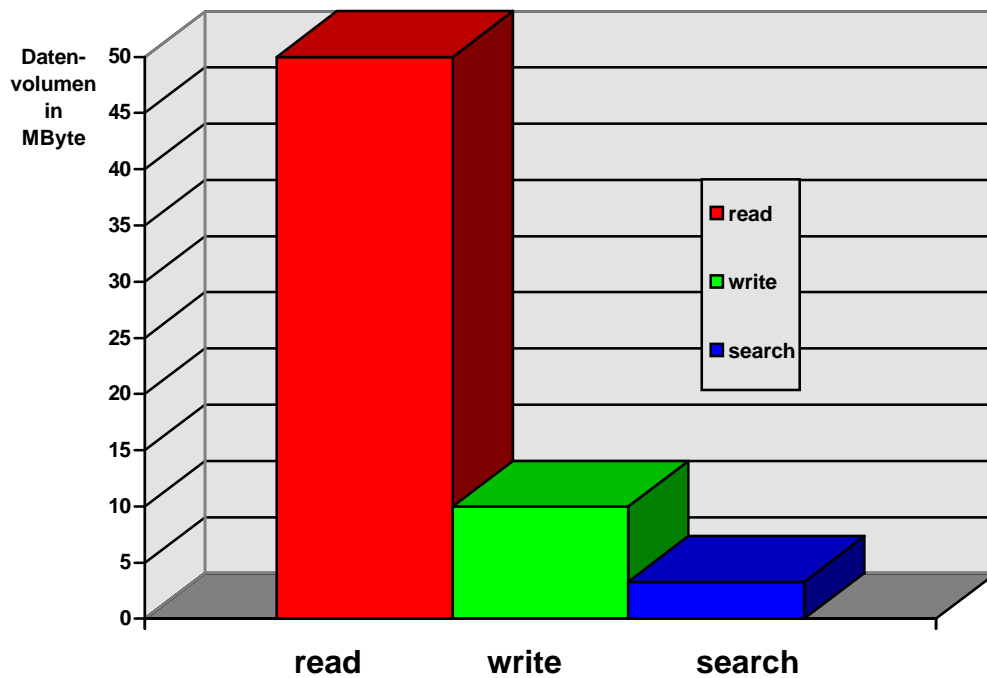
$$0.0018 / 7 = 0.025 \%$$

Fazit: Werden diese Werte überschritten, muß eine detaillierte Suche nach möglichen Ursachen sowohl auf der Netzwerk- als auch auf der Host-Seite erfolgen!

### ***NFS-RPCs und Datenvolumen während der Tests im Vergleich***



**Abbildung 23: Anzahl der NFS-RPCs der Testoperationen im Vergleich**



**Abbildung 24: bewegte Datenvolumen der Testoperationen im Vergleich**

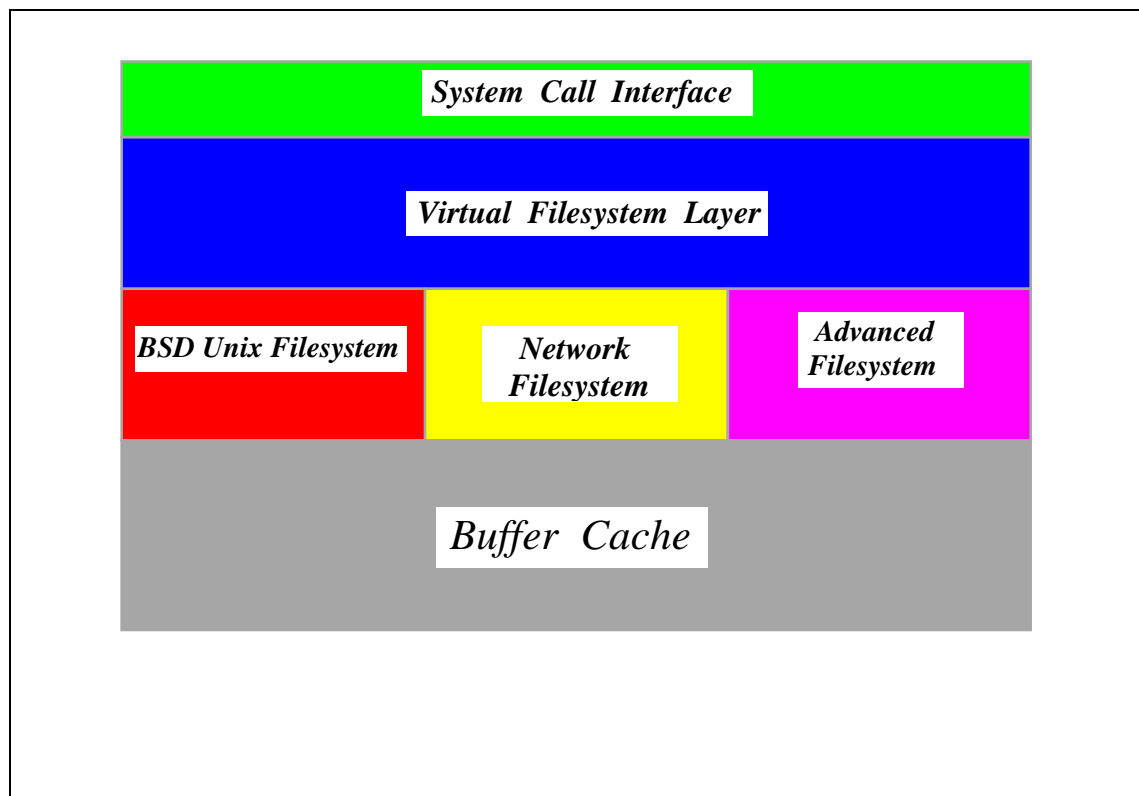
## Einfluß des Caches auf NFS-Operationen

Um den Einfluß von Dateisystem-Caches Rechnung zu tragen, werden in der folgenden Tabelle typische Antwortzeiten beim Lesen/Kopieren einer 2 MByte Datei von einer RZ58 Festplatte in den Hauptspeicher präsentiert.

<i>cp 2MB_file.dat /dev/null</i> (Antwortzeit)	lokaler I/O (DEC 3000-M400) (Sekunden)	NFS I/O (DEC 3000-M300) (Sekunden)
1. Versuch	0.71	2.01
2. Versuch	0.08	0.13
3. Versuch	0.08	0.13

Das Betriebssystem DEC OSF/1 optimiert Dateisystemzugriffe durch zwei Dateisystem-Caches. Der sogenannte *Traditional Unix Buffer Cache* nimmt beschreibende Datenstrukturen eines Dateisystems auf; diese sind beispielsweise *Inodes*, *Cylinder Group Blocks*, *Superblock* und *Directory Files*. Das Betriebssystem reserviert beim *Startup (Boot)* drei Prozent des physikalischen Hauptspeichers permanent für diesen Cache.

Die Datenblöcke werden in einem zweiten Cache, dem *Unified Buffer Cache*, zwischengespeichert. Die Größe dieses Caches variiert. Der Systemmanager kann jedoch obere und untere Schranken festlegen. Innerhalb dieser Werte konkurriert der *Unified Buffer Cache*, kurz *UBC*, mit dem *Virtual Memory Management* um freie Hauptspeicherseiten (*free list*). Je knapper die Hauptspeicherressourcen werden, desto aggressiver verhält sich das *Virtual Memory Management* gegenüber dem *UBC*, der dann bei der Standardeinstellung auf 10% der sogenannten *managed pages* reduziert wird. Auf einem System mit 64 MByte Hauptspeicher bleiben dann rund 5 MByte.



## *NFS-Statistiken aus dem Produktionsbetrieb*

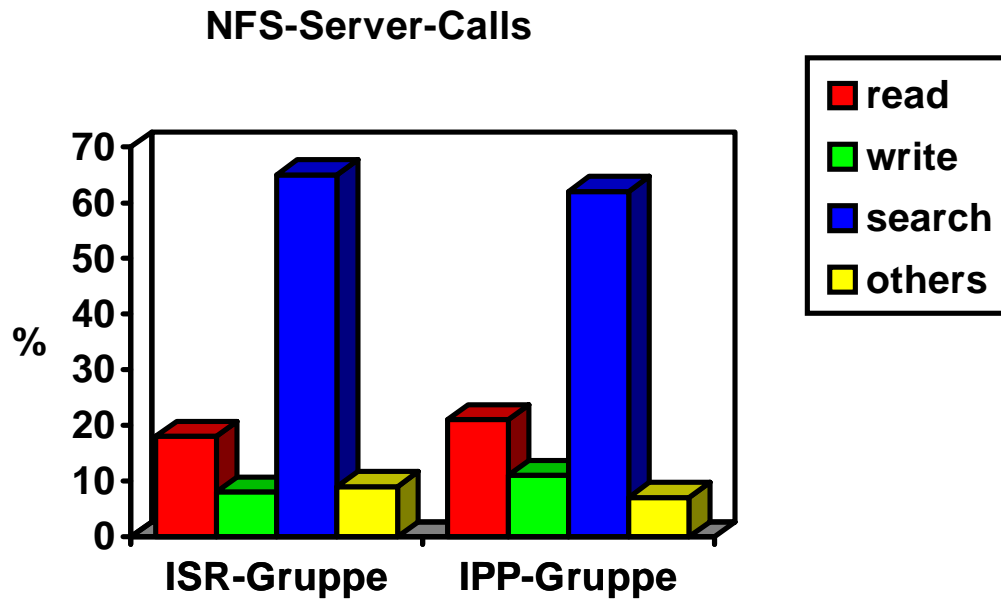


Abbildung 25: NFS-Server-Statistiken in AXP-Workstation-Gruppen

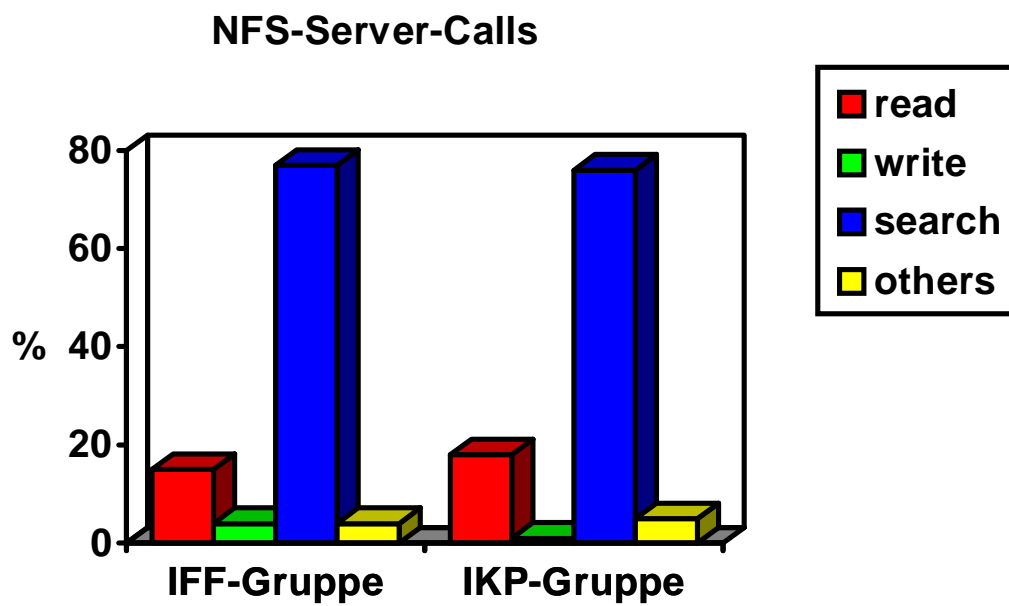
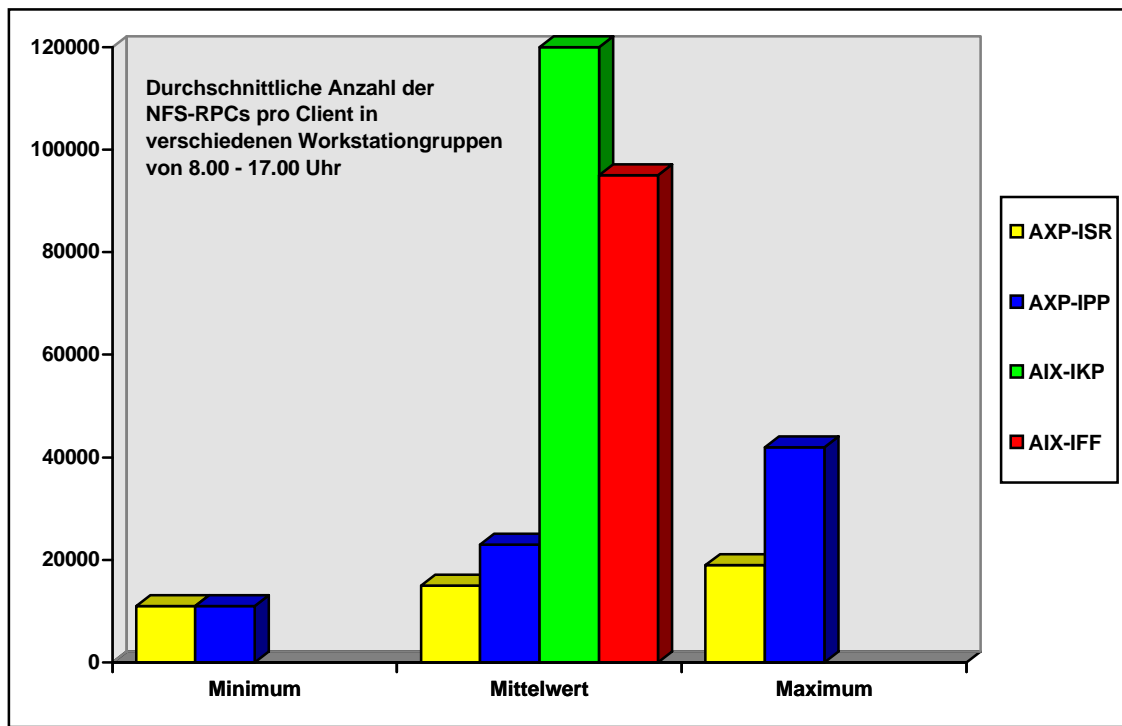


Abbildung 26: NFS-Server-Statistiken in AIX-Workstation-Gruppen



**Abbildung 27: Vergleich der Anzahl der RPC-Requests einer Client-Workstation in den verschiedenen Workstation-Gruppen**

Die Anzahl absoluter NFS-RPCs in den beiden betrachteten ALPHA-AXP-Workstation-Gruppen ist im Vergleich zu den AIX-Gruppen gering; durch die lokal installierten Betriebssystemkomponenten und die ebenfalls lokal installierten DEC-Layered-Products werden insbesondere die I/O-Anforderungen beim Starten von Anwendungen (*Stichwort: Demand Paged Executables*) von der lokalen SCSI-Peripherie bedient.

Die auf dem Server installierten Produkte, Public-Domain-Software und *nicht-DEC-Produkte*, nutzen fast alle *Shared Libraries*, die in der Regel von der lokalen Client-Festplatte geladen werden. In diesem Fall verteilen sich die I/O-Anforderungen auf das Netzwerksystem, also hier NFS, und das lokale Disksystem. Bei statisch gebundenen Anwendungen wie beispielsweise *Zmail* müssen dagegen bei der Aktivierung die I/O-Anforderungen überwiegend vom Netzwerksystem bedient werden.

Unabhängig von den hier präsentierten Werten und der im Fall der AXP-Gruppen abzulesenden momentan geringen Netzwerkbelastung muß festgehalten werden, daß abhängig von der Applikation (Beispiel: Bildverarbeitung) kurzfristig, im Falle eines Fehlers oder einer bestimmten Benutzeraktionen (z.B. Bearbeiten einer tar-Datei) sogar unmittelbar, eine völlig andere Situation entsteht, die die Unzulänglichkeiten einer Konfiguration basierend auf einem *Shared Ethernet Segment* aufdeckt. Der Vorteil einer Workstation-Gruppe liegt darin, dem jeweiligen Benutzer/Mitbenutzer einer Workstation in bestimmten Maßen den Zugriff auf Ressourcen wie CPU, Hauptspeicher und temporäre Speicherbereiche dauerhaft und mit einem hohen Grad an Robustheit zu garantieren, die nur schwer mit einer dem klassischen Time-Sharing äquivalenten Multiuser-Unix-Konfiguration erreichbar sind. Ein *Shared Ethernet Segment* kann heute aber von einer einzigen Workstation bzw. deren Benutzer monopolisiert werden und somit alle anderen Aktivitäten ernsthaft stören. Trotz stetig fortgeführter Strukturierungsmaßnahmen auf der Basis von Routern und Bridges wird diese Situation zunehmend zum Netzwerkalltag im KFA-net/INTERNET. Netzwerkmanagement muß in Folge neben den bisherigen Aufgaben Benutzeraktivitäten zeitlich koordinieren und Fehler aufgrund von Überlast analysieren. Die nötige Robustheit kann an dieser Stelle durch den Einsatz von ETHERNET/FDDI-Switches erreicht bzw. wieder hergestellt werden. Ferner ist der Gruppenserver so auszustatten, daß dieser auf jeden Fall mehrere voll aktive Client-Workstations bedienen kann.

Letztendlich ist die Beschaffung von Workstation-Gruppen und aktiven Netzwerkkomponenten gezielter aufeinander abzustimmen.

## Hintergrundinformationen zu SCSI

Bei einer FDDI-Verbindung zwischen File-Server und Workstations gewinnt eine optimale Konfiguration des I/O-Systems auf der Server-Seite an Bedeutung. Die hier in den Testsystemen verwendeten SCSI-Komponenten rangieren von den Leistungsdaten her, teilweise altersbedingt, am unteren Ende der Leistungsskala für SCSI II /Fast SCSI II - Geräte. Insbesondere SCSI-Komponenten von Fremdanbietern weisen oft höhere Leistungsdaten auf. Der Anschluß solcher Komponenten an ALPHA AXP Systeme hat sich in der Vergangenheit als unproblematisch erwiesen; um allerdings das Leistungspotential voll ausnutzen zu können, muß Feinarbeit geleistet werden. So steuert beispielsweise der SCSI-CAM Treiber der ALPHA AXP Systeme Fremdplatten nach Standard SCSI II Norm. Optionale Erweiterungen wie beispielsweise *Command Queuing* werden bei der Ansteuerung von Fremdplatten nicht verwendet. Optimale Ergebnisse werden erst nach einer Anpassung/Erweiterung des Treibers erzielt, durch die dann beispielsweise spezielle Fähigkeiten der Fremdplatte genutzt werden. Der File-Server sollte nach Möglichkeit mit mehreren FAST SCSI II Host Adaptern ausgestattet sein, um die Plattenzugriffe angemessen verteilen zu können. Werden per NFS Daten auf den File-Server geschrieben, ist ferner der Einbau eines NFS-Beschleunigers, Prestoserve, in Betracht zu ziehen.

Da die Plattenelektronik und -mechanik mittlerweile bereits von der Leistungsfähigkeit her die maximale Fast-SCSI II Transferrate erreicht oder beispielsweise im Fall von *Solid State Disks* sogar überschreitet, wird zunehmend auf der Hostseite Fast-Wide-SCSI, eine optionale Erweiterung von SCSI II, eingesetzt. Die Erweiterung des Bussystems von 8 Bit auf 16 Bit offeriert bei 10 MHz Transferleistungen von maximal 20 MByte/s. Dabei ist von Bedeutung, daß die SCSI-Spezifikation den Gemischtbetrieb alter 8-Bit-Geräte und neuer 16-Bit-Geräte vorsieht. Moderne Festplatten werden ebenfalls zunehmend mit Fast-Wide-SCSI-Interface ausgestattet.

<i>SCSI-Schnittstelle</i>	Daten-bits	max. Transferrate (MByte/s)	max. Anzahl Laufwerke	max. Bus-Länge (m)
single-ended	8	5	7	6
single-ended-fast	8	10	7	3
single-ended-wide	16	20	15	3
differential	8	5	7	25
differential-fast	8	10	7	25
differential-wide	16	20	15	25

**Tabelle 13: SCSI II - Bus - Spezifikationen**

Typ	Schnittstelle	Cache (KByte)	mittlere Positionier- zeit (ms)	Transferrate von/zum Medium (KByte/s)	zoned recording	Kapazität (MByte)
RZ58 *	single-ended	256	12.5	2500	+	1380
RZ26 *	single-ended	512	10	2600	-	1055
RZ26L	single-ended- fast	512	9.5	5500	+	1055
RZ28	single-ended- fast	1024	9.5	5500	+	2100
RZ26L- VW	single-ended- wide					
DSP3210	single-ended- fast	1024	6.3	5355	+	2050
DSP5400	single-ended- fast	512	7.0	5355	+	3815
ST12550N Barracuda	single-ended- fast	1024	5.1	5760	+	2040

\* == *nicht mehr im Angebot/Handel*

**Tabelle 14: Leistungsdaten aktueller SCSI-Laufwerke (Herstellerangaben)**



## Zusammenfassung

Beim Einsatz von Workstation-Gruppen mit leistungsfähigen Client- und Serverworkstations in einem NFS-Verbund sind, abhängig von der zur Verfügung stehenden Netzwerkinfrastruktur, deutlich unterschiedliche Datendurchsatzraten in den Gruppen festzustellen. Diese Unterschiede wirken sich teilweise bis zu den Antwort- und Programmladezeiten beim interaktiven Arbeiten innerhalb der Gruppe aus, weshalb der Auswahl und dem Ausbau der Netzwerkinfrastruktur eine wesentliche Bedeutung beim Erhalt und der Verbesserung der Leistungsfähigkeit von Workstation-Gruppen zukommt.

Die z.Zt. noch am häufigsten anzutreffende Netzwerkinfrastruktur basiert auf einer Verkabelung nach dem 10Base2-Standard ("ThinNet") und schließt eine große Anzahl von Netzwerkteilnehmern an ein flaches, unstrukturiertes Ethernet an. Der Nachteil dieser Topologie, die der Testkonfiguration 1 entspricht, liegt neben dem Low-Speed (10 MBit/s) Zugang zum Fileserver vor allem in der Konkurrenzsituation vieler Teilnehmer an einem Ethernet, die der Workstation-Gruppe zu keinem Zeitpunkt auch nur einen geringen Datendurchsatz innerhalb der Gruppe garantieren kann.

Eine Netzwerkinfrastruktur, bei der der jeweilige File-Server und die entsprechenden Workstations in einem durch eine Bridge abgetrennten Ethernet-Segment zusammengefaßt werden (Testkonfiguration 2), wird demgegenüber dem Kommunikationsprofil einer Workstation-Gruppe eher gerecht, da dem Client/Server-Verbund die volle Ethernet-Bandbreite von 1,25 MByte/s garantiert zur Verfügung steht und eine Konkurrenzsituation mit anderen Ethernet-Nutzern (Standalone-Geräte und auch weitere Workstation-Gruppen) vermieden wird. Diese Form der Strukturierung stellt gegenüber der Testkonfiguration 1 kaum größere Ansprüche an die Ausstattung der beteiligten Workstations, lediglich die Verkabelung muß eine Organisation der Workstations in einem eigenen Ethernet-Segment ermöglichen. Dies kann durch Benutzung des 10BaseT- bzw. 10BaseFB/FL-Standards mit einer sternförmigen Verkabelung auf UTP-, STP- oder Glasfaserbasis oder aber bei einer geeigneten räumlichen Verteilung der beteiligten Workstations auch mit der alten 10Base2-Verkabelung erreicht werden.

Eine Verbindung von Client- und Serverworkstations über einen Ethernet-FDDI-Switch, wie in der Testkonfiguration 3 vorgestellt, ist die Netzwerkinfrastruktur, die den heutigen Möglichkeiten der Client- und Serverworkstations in Punkto CPU-Leistung und Datendurchsatz für den Produktionseinsatz ein leistungsfähiges Netzwerk zur Verfügung stellt, das in Bezug auf den Gesamtdurchsatz durchaus noch Reserven für leistungsfähigere Server bereithält. Lediglich bei Betrachtung des zwischen einem Client und dem Server möglichen Maximaldurchsatzes ist diese Konfiguration nicht so leistungsfähig wie eine Netzwerkinfrastruktur ausschließlich auf FDDI-Basis, wie mit der Testkonfiguration 4 vorgestellt. Bezüglich der Verkabelungsvoraussetzungen unterscheidet sich die Konfiguration 3 gegenüber der Konfiguration 2 dadurch, daß eine sternförmige Verkabelung Voraussetzung zur Realisierung dieser Konfiguration ist und nicht durch optimale Standorte der Workstations und eine 10Base2-Verkabelung ersetzt werden kann.

Eine Variante der Netzwerkkonfiguration 3 ist in Abbildung 28 zu sehen. Durch den Einsatz eines zweiten FDDI-Interfaces am Switch kann ein FDDI-Backbone-Ring mit zentralen Servern (z.B. Backup- oder Archiv-Server) vom Workstation-Gruppen-Server durchgängig mit FDDI-Geschwindigkeit angesprochen werden.

Die Netzwerkkonfiguration 4, die eine Kopplung der Workstations über FDDI vorsieht, ist diejenige Lösung, die die Anforderung nach gutem Gesamtdurchsatz und Peakperformance für die einzelne Workstation am besten erfüllt. Ist evtl. bereits eine UTP oder S-UTP Verkabelung vorhanden und können damit die entsprechenden FDDI-Komponenten eingesetzt werden, so fallen auch die Kosten für die Realisierung der Netzwerkkonfiguration 4 kaum höher aus als die der Switchvariante. Zu beachten ist dabei allerdings, daß mit der FDDI-Lösung jeweils nur eine diskrete Anzahl von Workstations vom Einsatz des entsprechenden Equipments profitiert, während beim Einsatz eines Switches als Strukturkomponente mit z.B. 16 Ethernetports a 10 MBit/s ein Ethernet mit einer großen Anzahl von Teilnehmern in Abhängigkeit von der Kommunikationsstruktur z.B. über ein Hub-System so organisiert werden kann, daß ein theoretischer Gesamtdurchsatz des Systems von 160 MBit/s erreicht wird und jeder Netzwerkteilnehmer von dieser Strukturierung profitiert.

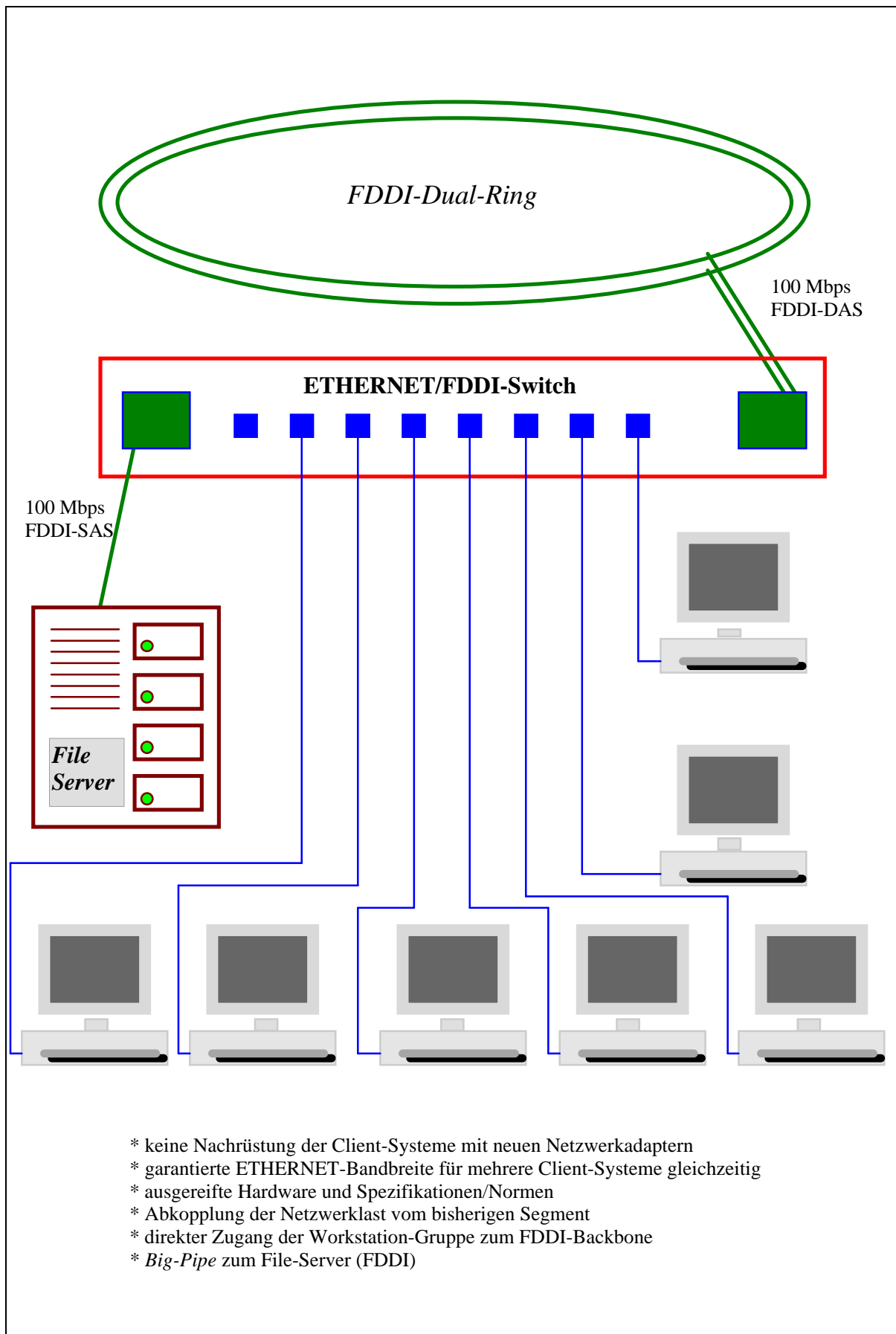


Abbildung 28: Ethernet-FDDI-Switch mit Server-FDDI- und Backbone-FDDI-Anschluß

## Literatur

- *Digital Industrial Networks Guidebook* (Digital Equipment)
- *Internetworking Technology and Acronyms* (Cisco Systems)
- *Digital Storage Technology Handbook* (Digital Equipment)
- *Gateway Ausgabe 6,7,8, 94 und 1,2 95* (Verlag Heise)
- *Guide to Prestoserve for DEC OSF/1* (Digital Equipment)
- *Fiber Distributed Data Interface* (Digital Equipment)
- *SCSI-Bus und IDE-Schnittstelle* (Addison-Wesley Verlag)
- *Interner Bericht Alpha AXP-Workstation-Gruppe* (KFA-Jülich, ZAM)
- *TCP/IP und NFS in Theorie und Praxis* (Addison-Wesley Verlag)
- *Galactica Network Switching Hub User's Guide* (Chipcom Corp.)